University of Alberta

NMR Structural Studies of the Human Rhinovirus 3C Protease

by

Trent C. Bjorndahl

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Faculty of Pharmacy and Pharmaceutical Sciences

Edmonton, Alberta
Spring, 2007

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-29653-0
Our file Notre référence
ISBN: 978-0-494-29653-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Dedication

This thesis is wholeheartedly dedicated to my wife Stephanie and my sons Nathan and Joshua. Their love, support and personal sacrifices over the past seven years is a major reason this document has come to fruition.

Abstract

The human rhinovirus (HRV) is a positive sense RNA virus responsible for about 30% of upper respiratory tract infections or "common colds". This virus relies on a 182 residue cysteine protease (3C gene product) to proteolytically processes its single gene product, to down regulate host cell processes and to bind to viral RNA (necessary for RNA transcription). Inhibition of this enzyme *in vitro* and *in vivo* has consistently demonstrated cessation of viral replication. This fact suggests 3C protease inhibitors could serve as good drug candidates. This thesis explores the structures of two states for the rhinovirus (serotype 14) 3C protease (apo and acetyl-LEALFQ-ethylpropionate inhibited) *via* nuclear magnetic resonance. The inhibited form allowed for a comprehensive analysis of the proteolytic pharmacophore. Furthermore, a comparison with the X-ray structure of the 3C protease from rhinovirus serotype 2 (51% sequence identity) bound to a peptidomimic inhibitor allowed the identification of serotype conserved intermolecular interactions involved in proximal substrate binding. In addition, the use of an extended peptidyl inhibitor permitted the study of downstream substrate interactions previously uncharacterized. Structural and dynamic comparisons between the two states showed only minor conformational changes upon inhibition. However, dynamic changes on the slow time scale were evident. The dynamics were characterized with $^{1}H/^{2}H$ exchange and showed that these differences were localized

within the enzyme's C-terminal β-barrel domain, which contains the proteolytic recognition site. The RNA binding site, which resides on the opposite side of the protease relative to the proteolytic site, remains structurally identical and presents an overall exchange rate in the apo form that is proportional to the N-terminal domain. This thesis presents the first solution structure and the first complete set of chemical shift data for any picornaviral 3C protease. These data can now facilitate the study of interactions between the 3C protease with ligands and other picornaviral or host cell proteins. These latter studies might help answer some biological questions, specifically, what other roles the 3C protease might play in the picornaviral life cycle.

Acknowledgments

I would like to thank the following people and organizations:

(i) The Health Research Foundation: HRF/Rx&D, The Protein Engineering Centre of Excellence (PENCE), the National Sciences and Engineering Research Council of Canada (NSERC) and the University of Alberta (Walter H. Johns Scholarship) for funding myself and this project.

(ii) Ryan McKay, Derek Webb and The Canadian National High Field NMR Centre (NANUC) for their assistance and use of the facilities.

(iv) Raj Jain, Carlos Valesques, Praveen Rao and Lena Andrew for help with the Acetyl-LEALFQ-ethylpropenoate inhibitor synthesis.

(iii) Jack Moore at the Alberta Peptide Institute (API) and Paul Semchuk at the Institute of Biomolecular Design (IBD) for running ES-MS and MALDI-TOF mass spectrometry on the Acetyl-LEALFQ-ethylpropenoate inhibitor and its intermediates and the labeled apo and inhibited HRV14-3C samples respectively.

(iv) Vishwanatha Somayaji for running 1D ^{1}H spectra of the Acetyl-LEALFQ-ethylpropenoate inhibitor and its intermediates.

(v) Michael N.G. James and Ken Ng for supplying the pET-3a plasmid containing the HRV14-3ABC gene and to Ping Li for performing the transfection into BL21(DE3) pLysS *E. coli* cells. Steve Bryson for insightful discussions regarding HRV14-3C protein purification and Valentyna Semenchenko for subsequent apo HRV14-3C protease production and NMR sample preparation.

(vi) David Wishart, Ed Knaus and Leo Spyracopoulos for all their support and guidance through my graduate career. An extra special thanks is extended to my supervisor and mentor, Dr. David Wishart, who had enough faith in my abilities and offered me a position in his lab after a ten year academic hiatus. Also for the many unprecedented opportunities that gave me a broad knowledge base beyond my expectations. Finally, for his wealth of expertise and input, which led to a successful completion of this project and for his financial and personal support throughout my studies.

(vii) My parents for a happy, good upbringing. For instilling confidence in myself and for the simple life lesson to "do what you enjoy". For with enjoyment comes dedication and only with dedication and hard work will result follow. Special thanks to all my family for their encouragement and support, which eased much of the burden imposed by graduate studies.

Contents

1 Introduction **1**

1.1 Upper Respiratory Tract Infections 1

1.1.1 Epidemiology and Etiology 1

1.1.2 Clinical Presentation 2

1.1.3 Physiology and Life-cycle of Rhinoviruses 2

1.1.4 Rhinovirus Therapy 6

1.2 NMR Theoretical Principles 10

1.2.1 Chemical Shifts 14

1.2.2 Coupling Constants 15

1.2.3 Nuclear Overhauser Effect 18

1.2.4 Physical Parameter Considerations 19

1.3 Macromolecular Structure Determination 20

1.3.1 NMR Sample Preparation 21

1.3.2 Data Collection and Sequential Assignment 22

1.3.3 Structure Calculation 25

1.4 Thesis Outline 27

2 Sample Preparation **35**

2.1 Introduction 35

2.2 Materials and Methods 36

2.2.1 SDS-PAGE Gel Analysis 37

2.2.2 Enzymatic Activity Assay 37

2.2.3 Bradford Assay 37

2.2.4 Chromatography Resin Binding Experiments 38

2.2.5 HRV14-3C Protease Expression 38

2.2.6 HRV14-3C Protease Purification 39

2.2.7 Button Test 40

2.2.8 Refolding Procedure 41

2.2.9 NMR Sample Preparation 42

2.3 Results and Discussion 42

2.3.1 HRV14-3C Protease Expression 42

2.3.2 HRV14-3C Protease Purification 44

2.3.3 HRV14-3C Chromatography Fraction Detection 46

2.3.4 Protease Quantification 47

2.3.5 NMR Sample Stability 48

2.4 Conclusion 49

3 Inhibitor Synthesis **53**

3.1 Introduction 53

3.2 Materials and Methods 55

3.2.1 Inhibitor Synthesis 55

3.2.2 Mass Spectrometry 58
3.2.3 NMR 59
3.3 Results and Discussion 60
3.3.1 Inhibitor Synthesis 60
3.3.2 Inhibitor Solubility and Enzyme Activity Assays 64
3.4 Conclusion 66

4 Inhibited HRV14-3C Structure **71**

4.1 Introduction 71
4.2 Materials and Methods 72
4.2.1 Sample Preparation 72
4.2.2 NMR Experiments 72
4.2.3 NMR Chemical Shift and NOE Restraint Assignment 73
4.3 Results 74
4.3.1 Inhibited HRV14-3C Chemical Shift Assignments 74
4.3.2 Inhibited HRV14-3C NMR Structure Calculation 79
4.4 Discussion 83
4.4.1 The Inhibited HRV14-3C Protease Structure 83
4.4.2 Induced Fit Substrate Recognition 85
4.4.3 Active Site Triad Comparison 87
4.4.4 Substrate Binding: Backbone and Sidechain Contacts 91
4.4.5 Substrate Specificity and Pharmacophore Analysis 93
4.5 Conclusion 101

5 Apo HRV14-3C Structure **108**

5.1 Introduction 108

5.2 Methods and Materials 108

5.2.1 NMR Sample Preparation 108

5.2.2 NMR Data Collection 109

5.2.3 Correlation Time Calculation 111

5.2.4 Hydrogen/Deuterium Exchange Rate Analysis 112

5.3 Results 112

5.3.1 Apo HRV14-3C Chemical Shift Assignments 112

5.3.2 Apo NMR HRV14-3C Structure Calculation 113

5.4 Discussion 119

5.4.1 Apo vs. Inhibited HRV14-3C Protease Chemical Shift Assignments . 119

5.4.2 Apo vs. Inhibited HRV14-3C Protease Solution Structure 120

5.4.3 Deuterium Exchange and HRV14-3C Dynamics 125

5.4.4 Oligomerization Analysis 130

5.5 Conclusion 133

6 Conclusion **137**

6.1 Introduction 137

6.2 Biological Questions 138

6.2.1 Heterodimerization Interface Identification 138

6.3 Inhibitor Design 139

6.3.1 Proteolytic Inhibitors 139

6.3.2 Inhibition with Zinc 140

6.3.3 Other Methonds of Inhibition 140

6.4 Conclusion 142

A Molecular Biology Recipes **145**

B Expression and Purification Protocols **150**

B.1 Expression Protocol 150

B.2 Protein Induction Test 151

B.3 Protein Purification 152

C XPLOR / CNS input scripts **153**

C.1 Ethylpropionate *topology* File 154

C.2 Ethylpropionate *parameter* File 156

C.3 XPLOR/CNS *generate.inp* File Modifications 157

C.4 Peptide Linkage File 158

C.5 RECOORD Input File Modifications 159

D Chemical Shift Assignments **160**

D.1 Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{13}C Chemical Shifts 160

D.2 Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{1}H Chemical Shifts 161

D.3 Apo HRV14-3C Chemical Shifts 162

D.4 Inhibited HRV14-3C Chemical Shifts 167

E Protection Factor Calculations **173**

List of Tables

1.1 T_1, T_2 and τ_c vs. Protein Molecular Weight 19

2.1 Button Test Results 48

4.1 NMR Experiments Conducted on the Inhibited HRV14-3C Protease 73

4.2 Structural Statistics for the Inhibited HRV14-3C Protease 82

4.3 Backbone Angles of the Picornaviridae $3C^{pro}$ II_e - II_f Loop/Turn Region . 100

5.1 Backbone NMR Experiments for the Apo HRV14-3C Protease 110

5.2 Structural Statistics for the HRV14-3C Proteases 118

5.3 Superposition of Picornaviridae $3C^{pro}$ RNA Binding Region 123

5.4 Apo and Inhibited HRV14-3C K_{ex} and P_{factor} Data 129

D.1 Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{13}C Chemical Shift Assignments 160

D.2 Acetyl-LEALFQ-ethylpropenoate Inhibitor ^{1}H Chemical Shift Assignments 161

D.3 Apo HRV14-3C Chemical Shift Assignments 162

D.4 Inhibited HRV14-3C Chemical Shift Assignments 167

E.1 Apo HRV14-3C K_{rc} Calculations 173

E.2 Inhibited HRV14-3C K_{rc} Calculations 174

E.3 Apo HRV14-3C K_{ex} and P_{factor} Calculations 176

E.4 Inhibited HRV14-3C K_{ex} and P_{factor} Calculations 177

List of Figures

1.1 Picornaviridae Gene Product 3

1.2 Picornaviridae Life Cycle 5

1.3 Cysteine Protease Catalytic Mechanism 8

1.4 3C Protease Inhibition with an Ethylpropenoate Michael Acceptor 9

1.5 1D NMR Experiment 13

1.6 1D ^{1}H NMR Spectrum of Ethanol 15

1.7 J-Coupling Constants 16

1.8 $^{3}J_{HNH\alpha}$ Karplus Equation Relationship 17

1.9 β-Strand Secondary Structure 18

1.10 Protein τ_c vs. Spectral Linewidths 20

1.11 Multiple Dimension NMR Experiments 23

1.12 Sequential Connectivity 24

1.13 NOE Data Derivation 26

1.14 NOEs for Val^{47} 27

2.1 DEAE-Sepharose™ Chromatography Results 41

2.2 SDS-PAGE Gel: Induction Experiments 43

2.3 Sepharose™ Chromatography Resin Binding Experiment 44

2.4 Q-Sepharose™ Purification of the HRV14-3C Protease 45

2.5 $UV_{200-360nm}$ Scan of the HRV14-3C Protease 46

2.6 Preliminary 1D ^{1}H NMR Spectrum of the HRV14-3C Protease 49

3.1 HRV2-3C vs. Polio-3C B-factors 54

3.2 Structural Relationship of Flexible Residues in Picornaviridae 3C Proteases 55

3.3 Inhibitor Synthesis 56

3.4 ES-MS Spectrum of the Acetyl-LEALFQ-Ethylpropenoate Inhibitor 59

3.5 MALDI-TOF Spectra of the [U-$^{13}C/^{15}N$]-HRV14-3C Samples 60

3.6 Chemical Structure and 1D ^{1}H NMR Spectrum of the Acetyl-LEALFQ-ethyl Propenoate Inhibitor in DMSO 61

3.7 ^{1}H-TOCSY of the Acetyl-LEALFQ-ethyl Propenoate Inhibitor in DMSO . 62

3.8 Cis/Trans Diastereoisomers of the Ethyl Propenoate Group 63

3.9 ^{13}C-HSQC of the Acetyl-LEALFQ-ethyl propenoate Inhibitor in DMSO . . 65

3.10 HRV14-3C p-Nitroaniline Colorimetric Assays 67

4.1 ^{15}N-HSQC Spectra for the Inhibited HRV14-3C Protease 75

4.2 HNCACB Strip Plots from the Inhibited HRV14-3C 76

4.3 C(CO)NNH Strip Plots from the Inhibited HRV14-3C 77

4.4 HNCACB Strip Plots for Region QKIRVKDK$^{52-59}$ of the Inhibited HRV14-3C Protease 78

4.5 CCH-TOCSY and HCCH-TOCSY Strip Plots for the Leu^{123} 79

4.6 A Portion of the ^{13}C-edited NOESY-HSQC Spectrum for the Inhibited HRV14-3C Enzyme 80

4.7 Ribbon Representation of the Inhibited HRV14-3C Protease 84

4.8 ^{13}C -NOESY-HSQC Spectra for His^{160}'s $^{1}H_{\delta 2}$. 85

4.9 Chemical Shift and Solvent Accessible Surface Area Changes 87

4.10 Hydrogen Bonds Within the C-terminal β-barrel Domain 88

4.11 Electrostatic Molecular Surface of the HRV14-3C and HRV2-3C Proteases . 89

4.12 Catalytic Triad and P_1 to P_3 Substrate Interactions 91

4.13 Schematic Representation of the Pharmacophore Interactions 92

4.14 P_1 and P_2 Substrate Interactions 94

4.15 P_3 Substrate Interactions 95

4.16 P_4 to P_6 Substrate Interactions 96

4.17 Substrate Interactions with HRV14-3C's Asn^{164} Residue 98

4.18 Substrate Interactions with HRV2-3C's Asn^{165} Residue 99

5.1 ^{15}N-HSQC Spectra of the Apo HRV14-3C Protease 114

5.2 HNCACB Strip Plots from the Apo HRV14-3C Protease 115

5.3 C(CO)NNH Strip Plots from the Apo HRV14-3C Protease 116

5.4 ^{13}C -NOESY-HSQC Spectrum for Ile^{68} $^{1}H_{\gamma 2}$ of the Apo HRV14-3C Protease 117

5.5 Suspected Residues Undergoing μs Conformational Exchange 121

5.6 Normalized Atomic Displacements and Per-residue NOE Restraint Count for the Apo and Inhibited HRV14-3C Proteases 122

5.7 Asn^{14} Side-chain Hydrogen Bonding in the Inhibited HRV14-3C Protease . 124

5.8 HNHA Strip Plots for the Apo and Inhibited HRV14-3C Protease 125

5.9 Apo vs. Bound P_{factor} Data 126

5.10 Deuterium Exchange ^{15}N-HSQC Spectra of the Inhibited HRV14-3C Protease 127

5.11 Deuterium Exchange ^{15}N-HSQC Spectra of the Apo HRV14-3C 128

5.12 K_{ex} Data for the HRV14-3C Protease 131

6.1 Picornaviridae 3C Protease Sequential Alignment 141

List of Abbreviations

APS	-	Ammonium Persulfate
BB	-	Backbone
BMRB	-	BioMagResBank
BSA	-	Bovine Serum Albumin
CM	-	Carboxymethylcellulose
COPD	-	Chronic Obstructive Pulmonary Disease
CSI	-	Chemical Shift Index
dd	-	Double Distilled
DCM	-	Dichloromethane
DEAE	-	Diethylaminoethyl
DIBAL	-	Diisobutylaluminum Hydride
DMF	-	Dimethylformamide
DMSO	-	Dimethylsulfoxide
DTT	-	Dithiothreitol
EDC	-	1-(3-(Dimethylamino)propyl)-3-ethylcarbodiimide Hydrochloride
EDTA	-	Ethylenediaminetetraacetic Acid
EM	-	Electromagnetic

LIST OF ABBREVIATIONS

EtOAc	-	Ethyl Acetate
FID	-	Free Induction Decay
FMD	-	Foot-and-Mouth Disease
HAV	-	Hepatitis-A Virus
HBTU	-	O-Benzotriazole-N,N,N′,N′-tetramethyl-uronium-hexafluoro-phosphate
HOBt	-	Hydroxybenzotriazole Hydrate
HRV	-	Human Rhinovirus
HRV14	-	Human Rhinovirus (Serotype 14)
HRV14-3C	-	Human Rhinovirus (Serotype 14) 3C Protease
HRV2-3C	-	Human Rhinovirus (Serotype 2) 3C Protease
IBD	-	Institute for Biomolecular Design
ICAM	-	Intracellular Adhesion Molecule
IPTG	-	Isopropyl β-D-thiogalactopyranoside
LB	-	Luria Broth
MALDI-TOF	-	Matrix Assisted Laser Desorption - Time of Flight
MW	-	Molecular Weight
NMR	-	Nuclear Magnetic Resonance
NOE	-	Nuclear Overhauser Effect
NOESY	-	Nuclear Overhauser Effect Spectroscopy
NMM	-	N-Methylmorpholine
OD_{600}	-	Optical Density at 600 nm
PBS	-	Phosphate Buffered Saline
PCR	-	Polymerase Chain Reaction

LIST OF ABBREVIATIONS

PDB	-	Protein Data Bank
PEE	-	Polyethylenimine
PFG	-	Pulse Field Gradient
Q	-	Quaternary Ammonium
RMSD	-	Root Mean Squared Deviation
RPM	-	Revolutions Per Minute
SAR	-	Structure Activity Relationship
SDS	-	Sodium Dodecyl Sulfate
SDS-PAGE	-	Sodium Dodecyl Sulfate - Poly Acrylamide Gel
SP	-	Sulfopropyl
TB	-	Terrific Broth
TEMED	-	N,N,N′,N′-Tetramethylethylenediamine
TFA	-	Trifluoroacetic Acid
TFE	-	Trifluoroethanol
THF	-	Tetrahydrofuran
TIS	-	Triisopropylsilane
TLC	-	Thin Layer Chromatography
TMS	-	Tetramethylsilane
TOCSY	-	Total Correlated Spectroscopy
TRIS	-	2-amino-2-(hydroxymethyl)-1,3-propanediol
UV	-	Ultraviolet
U-^{13}C/^{15}N	-	Uniformly ^{13}C and ^{15}N isotopically labeled

List of Software Programs

ACD v9.0
NMR spectra processing and analysis
`http://www.acdlabs.com`

AQUA v3.2
NOE restraint plotting and violation computation (Linux)
`http://tang.bmrb.wisc.edu/~jurgen/aqua/`

CHEMSILICO
Inhibitor physical parameter predictions (online)
`http://chemsilico.com`

CorelDraw v12.0
Graphics editing and figure preparation (Windows)
`http://www.corel.com`

CYANA v2.1
Simulated annealing and energy minimization (Linux)
`http://www.las.jp/prod/cyana/eg/`

CNS v1.1
Simulated annealing and energy minimization (Linux)
`http://cns.csb.yale.edu/v1.1/`

The GIMP v2.2.12
Graphics editing and figure preparation (Linux)
`http://www.gimp.org`

GNUPLOT v4.0
Interactive data and function plotting tool (Linux)
`http://www.gnuplot.info/`

GRACE v5.1.20
WYSIWYG 2D plotting tool (Linux)
`http://plasma-gate.weizmann.ac.il/Grace/`

KILE v1.9.1
KDE Integrated LaTeX Environment (Linux)
`http://kile.sourceforge.net`

LIGPLOT v4.4.2
Ligand-enzyme intermolecular interaction representation (Linux)
`http://www.biochem.ucl.ac.uk/bsm/ligplot/ligplot.html`

NMRDraw v2.3
Companion graphical interface for NMRPipe (Linux)
`http://spin.niddk.nih.gov/bax/software/NMRPipe/`

NMRPipe v2.3
NMR spectrum processing (Linux)
`http://spin.niddk.nih.gov/bax/software/NMRPipe/`

NMRView v5.2.2
NMR spectrum visualization and assignment (Linux)
`http://onemoonscientific.com/nmrview/`

ORIGION v7.5
Mathamatical graphing and least squares curve fitting (Windows)
`http://www.originlab.com`

Procheck-NMR v3.5.4
NMR ensemble quality analysis (Linux)
`http://www.biochem.ucl.ac.uk/~roman/procheck_nmr/procheck_nmr.html`

P*y*MOL v0.98
PDB file visualization and structure rendering (Linux)
`http://pymol.sourceforge.net/`

RECOORD v1.0
Protein water refinement and energy minimization (Linux)
`http://www.ebi.ac.uk/msd-srv/docs/NMR/recoord/main.html`

STC v5.0
Free energy and surface area calculations for protein-ligand interactions (Linux)
`http://www.bionmr.ualberta.ca/bds/software/stc/latest/index.html`

VNMR v3.1b, v 3.1c
NMR experimental acquisition (Sun)
`http://www.varianinc.com`

XPLOR v3.851
Free energy calculations for protein-ligand interactions (Linux)
`http://atb.csb.yale.edu/xplor/`

XPLOR-NIH v2.10
Simulated annealing and energy minimization (Linux)
`http://nmr.cit.nih.gov/xplor-nih/`

Chapter 1

Introduction: Picornaviridae Pathology and NMR Methodology

1.1 Upper Respiratory Tract Infections

1.1.1 Epidemiology and Etiology

In 2004, the National Institutes of Health estimated that the 'common cold' accounted for ~1 billion illnesses in the United States alone [1]. The incidence breakdown is age dependent in which childred under 16 years of age acquire 6 - 8 infections, individuals aged 16 to 45 get 2 - 3 infections and adults over 45 years of age suffer from ~1 infection per year [2]. A number of epidemiological studies have reported various statistics. This variability is dependent on the demographic area being studied, the circulating virus at the time of study and the method of detection. A study conducted in 1997 reported ~50% of colds were caused by rhinoviruses [3]. A more recent study has confirmed that the human rhinovirus (HRV) is actually responsible for about 25 to 35% of all adult colds [4]. The genetically related enteroviruses account for a large portion of the remaining infections, followed by a variety of other organisms including, coronaviruses, adenoviruses and respiratory syncytial viruses [5], while 'cold' infections due to bateria are rare. With the recent advent of optical thin film and multiplex RT-PCR based detection methods for clinical use [6, 7, 8, 9], these numbers should become increasingly accurate.

1.1.2 Clinical Presentation

Rhinoviruses and enteroviruses have a tropism for the upper respiratory tract. This is maintained by their viral capsid, which dissociates at low pH and their preference to replicate at lower temperature (33 °C). These viruses exploit a number of cellular surface proteins as receptors. The rhinoviruses in particular utilize the ICAM-1 and LDLR receptors to gain entry into the host cell. Once viral inoculation of the nose or upper respiratory tract occurs, an incubation period of 12 to 72 hours (average 8 - 16 hours) ensues. During this period, the concentration of virons is greatest and persons are most contagious. The cold virus readily transfers *via* hand to hand [10] and hand to surface [11] contact. Clinical symptoms generally start with a sore throat, followed by nasal congestion, rhinorhea (runny nose) and sneezing. These symptoms increase in intensity over a 2 to 3 day period and can lead to other complications such as headache and loss of taste and smell. Cough occurs in ~30% of patients. These symptoms persist for generally 7 to 10 days, however, they can persist for up to two weeks in ~25% of patients. During this time, sleep disturbances can occur for about 4 nights. In patients with concomitant lower respiratory tract illnesses such as COPD and asthma, rhinovirus infections can either precipitate exacerbations [12] or become more virulent themselves [2]. Futhermore, rhinovirus infections have been linked to childhood wheezing and are the most significant risk factor among children suceptible to developing allergies and asthma [13].

These facts reiterate the fact that HRV infections are the most abundant viral infection worldwide and one of the leading causes of human disease and morbidity. Although the disease rarely results in mortality, the estimated economic impact in the United States alone is estimated to be 40 billion dollars annually. Each year, the impact of these illnesses translates into an estimated 90 million days of restricted activity and 45 million lost school days.

1.1.3 Physiology and Life-cycle of Rhinoviruses

Picornaviral Gene Product and RNA Translation

The Picornavirus (meaning 'small' RNA) family includes rhinoviruses and enteroviruses. As mentioned, members within this viral family exploit a number of host cell receptors

for gaining access into a host cell. For example, enteroviruses bind to the host's CD155, decay accelerating factor (DAF) and coxsackievirus and adenovrius receptor (CAR), while rhinoviruses have been shown to bind to the intracellular adhesion molecule (ICAM-1) and low density lipoprotein (LDL) receptors. Once a picornavirus viron particle gains entry into the host cell its positive sense RNA is released into the cytoplasm. The RNA genome is efficient at 'hijacking' the host cell replication machinery (ribosome and initiation factors) to process their 7.0-8.5 kb single stranded positive sense RNA genome (Figure 1.1) into a single 250 kDa gene product or polyprotein that contains all the structural and non-structural proteins necessary for viral replication. The polyprotein is divided into three distinct regions, P_1, P_2 and P_3 (Figure 1.1). The P_1 region contains the precursor viral capsid proteins while the P_2 and P_3 regions contain the precursors of functional proteins important for viral RNA replication. A number of the non-structural proteins are quick at halting host cell processes, thereby turning the host cell into a viral replication factory.

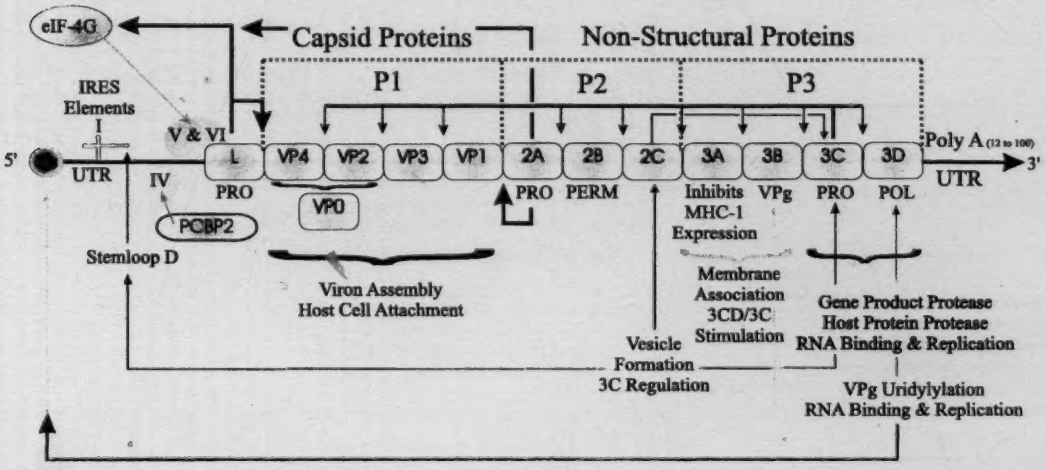

Figure 1.1: **Picornaviridae Gene Product.** The non-structural gene products (P_2 and P_3) have their corresponding functions listed (PRO=protease; POL=polymerase; PERM=cell permeability enhancer). The L and 2A proteases have their cleavage sites indicated with red arrows. Cleave sites along the picornaviridae gene product that are mediated by the 3C gene product are colored magenta. The 'L' protease is not present in the rhinovirus gene.

The viron gene contains a 5′ internal ribosome entry site (IRES) that contains a number of untranslated structured RNA elements (UTR), which serve as a site for both viral and

host cell protein binding. This binding is necessary for RNA translation and transcription. Rhinovirus RNA translation relies solely on the host cell's translational machinery including the cap-binding eukaryotic initiation factor (eIF-4G) and the poly(rC) binding protein (PCBP). Once the polyprotein is translated, a number of cleavage events in *cis* and in *trans* occur to separate the viral proteins from the polyprotein. These are mainly carried out by two virally encoded proteases identified as HRV-2A and HRV-3C. The primary cleavage is carried out co-translationally by the 2A protease, which separates the P_1 region from the P_2/P_3 region. Subsequently, the majority of the proteolytic processing within P_1, P_2 and P_3 is carried out by the 3C protease.

Host Cell Down-regulation

Once the polyprotein has been processed, a number of the functional proteins quickly shut down host cell activities. The 2A (and L protease in aphthoviruses) cleave host cell eukaryotic initiation factor (eIF-4G), thereby shutting down host cell translation. However, the RNA binding remnant of this protein is salvaged by the viral translation complex [14] and now has preferred binding to the viral IRES site. The viral 3D gene product contains a nuclear localization sequence (NLS), which is used to transport the fused 3CD gene product into the host cell nucleus [15]. Once inside the nucleus, the 3C protease self-cleaves from the 3CD heterodimer and proceeds to cleave the host cell RNA polymerase I, II and III, poly(A)-binding protein and transcription factors [16] between Gln-Gly sequence pairs. It has also been shown that the 2C polypeptide regulates the activity of the 3C protease [17]. In addition to the host protein cleavage carried out by the L, 2A and 3C proteases, cell membrane permeability is increased by the 2B protein [18] and the cellular secretory pathways are shut off by the 3A protein [19]. This latter event reduces the expression of MHC-1 molecules on the cell surface [20] thereby reducing the host's immune response against the infected cell.

RNA Transcription and Viron Assembly

In picornaviruses, the positive sense RNA genome is used to synthesize a negative sense RNA template, which in turn is used to replicate more positive sense RNA genomes. These genomes can be used for additional RNA translation, can template more negative sense

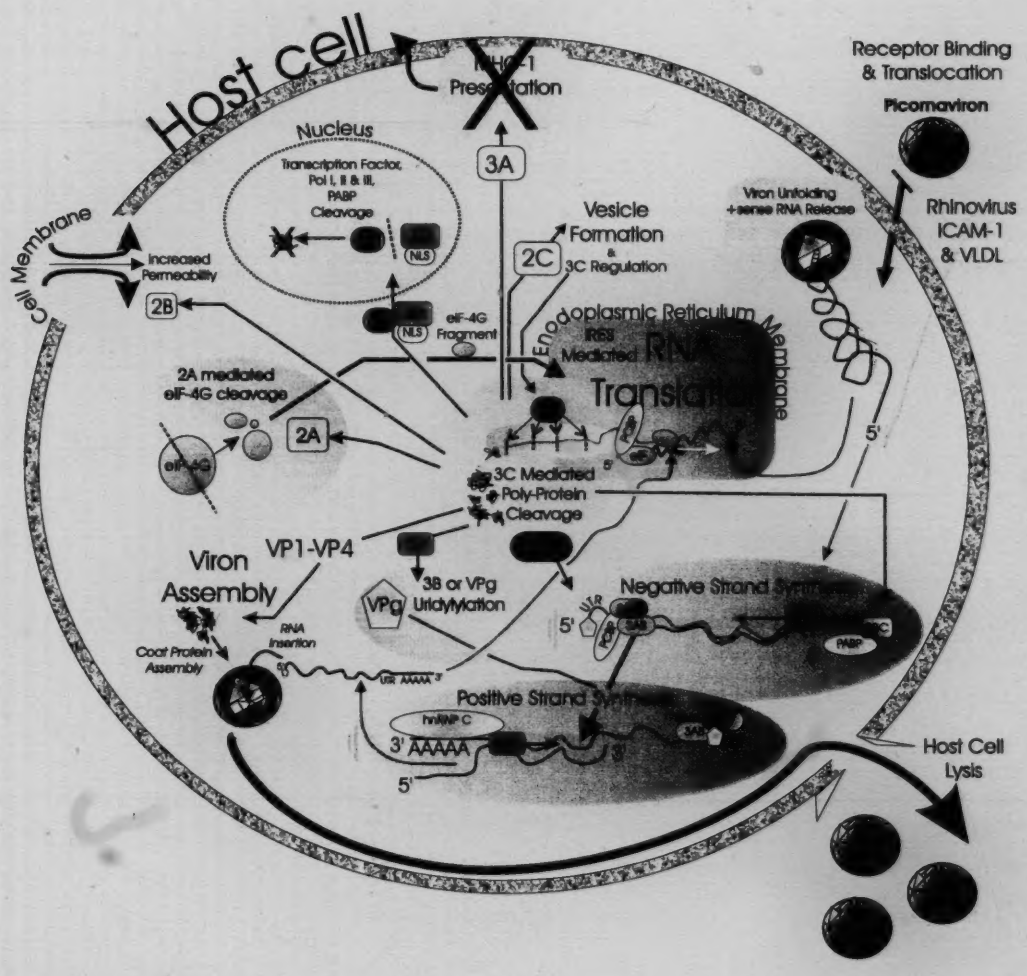

Figure 1.2: **Picornaviridae Life Cycle**. The life cycle for the picornavirus is generalized. Both the viral and host proteins involved with viral RNA translation and transcription are illustrated. NLS=nuclear localization sequence. PCBP=Poly r(C) binding protein. PABP=Poly (A) binding protein. hnRNP C=Heterogeneous nuclear ribonucleoprotein C.

RNA strand synthesis or they can be incorporated into newly formed viral capsids ready to infect more host cells.

The current opinion is that communication exists between the 5′ and 3′ ends of the viral RNA to mediate efficient RNA replication. As mentioned earlier, the 5′ UTR region of the viral genome contains structured RNA elements (IRES) necessary for RNA transcription as well as translation. These elements (identified as I to VI) are responsible for binding a number of proteins. Elements IV, V and VI bind the PCBP-2 and eIF proteins necessary for RNA translation. Upstream from these is the cloverleaf IRES I element known to bind the 3CD protein [21], which includes the viral RNA polymerase (3D gene product). The 3AB gene product binds both the 5′ and 3′ ends of the RNA [22]. It is also required for membrane association [23] and stimulation of the 3D polymerase [24].

The 3′ end contains a poly-adenine sequence (n=12 to 100), which binds host cell poly(A) binding protein (PABP). It also contains an UTR region that binds viral 2BC and 3D proteins. In particular, the 2C protein (in addition to regulating the 3C protease and playing a role in vesicle formation) has confirmed ATPase and GTPase activity, although it has not yet demonstrated helicase activity. Host cell proteins are thought to form initiation complexes with the viral RNA. One host protein, hnRNP C, has been clearly identified to bind to the 3′ end of negative sense viral RNA [25]. This protein along with a number of other host cell nuclear proteins have shown cellular rearrangement during picornaviral infection [26].

1.1.4 Rhinovirus Therapy

Rhinoviruses include more than 110 known serotypes that are grouped into two subgenuses, A and B. Rhinoviruses are composed of a capsid that contains four viral proteins VP_1, VP_2, VP_3 and VP_4, which are arranged into 60 repeating protameric icosahedral units. Genetic divergence within these capsid proteins are thought to be the cause of the antigenic diversity associated with these viruses. Such variability accounts for the diversity of cold symptoms, the frequency of colds and why incidence rates reduces with age. This kind of diversity, however, has hindered the development of a single cold vaccine [27, 28]. One capsid binding drug, Plecoranil, binds to the VP1 protein and is believed to stabilize capsid unfolding and interfere with receptor binding. It has shown to inhibit ~92% of rhinovirus

serotypes. However, because the incubation period of rhinoviruses is relatively short and the onset of clinical symptoms follow incubation, studies with Plecoranil have shown only a minor reduction of symptoms by ~1 day [29]. Another target for drug therapy is the 3C gene product, a relatively small 21 kDa cysteine protease. As discussed previously, the 3C protease is not only responsible for viral maturation and poly-protein processing, but also for RNA binding [30] and host-cell disabling through selective proteolytic cleavage of key host proteins [31]. This is graphically illustrated in Figure 1.2. Because HRV-3C protease plays an important role in viral maturation and inhibition of this enzyme has shown to be effective at halting viral replication [32], 3C protease inhibitors have been investigated as potential pharmaceutical agents in halting or treating some colds.

3C Protease Targeting

Despite significant efforts over the last 20 years to develop and structurally characterize inhibitors for the 3C protease, only three X-ray crystal structures of the HRV-3C protease have been reported [32, 33, 34, 35] from the 110+ available serotypes (these representatives include serotypes 2, 14 and 16). Furthermore, only one coordinate data set for an inhibitor-bound form of the HRV2-3C protease (PDB code 1CQQ) has ever been publicly released. These structural studies have shown that this 182-residue cysteine protease is composed of a two-domain, β-barrel structure similar to chymotrypsin-like serine proteases [33, 32, 34]. The active site is situated between the two domains and consists of a cysteine-histidine-glutamate catalytic triad along with an electrophilic oxyanion hole. The proposed mechanism of catalysis is outlined in Figure 1.3. A long, shallow grove in the C-terminal β-barrel domain accommodates the wide range of peptide substrates recognized along the polyprotein gene product [36, 37]. Interestingly, cysteine proteases, unlike serine proteases, tend to recognize 4-5 residue cleavage sequences rather than single residue cleavage points (lysine/arginine for trypsin, hydrophobic residues for chymotrypsin). By convention, these substrate sequences are numbered as P_n for residues preceding the scissile bond and P_n' for residues following the cleavage site. Residues from the P_5 (the fifth residue preceding the cleavage site) to P_3' (the third residue following the cleavage site) positions largely define the picornaviral 3C cleavage sequence.

Studies by Cordingley *et al.* [36] initially determined a minimal substrate recognition

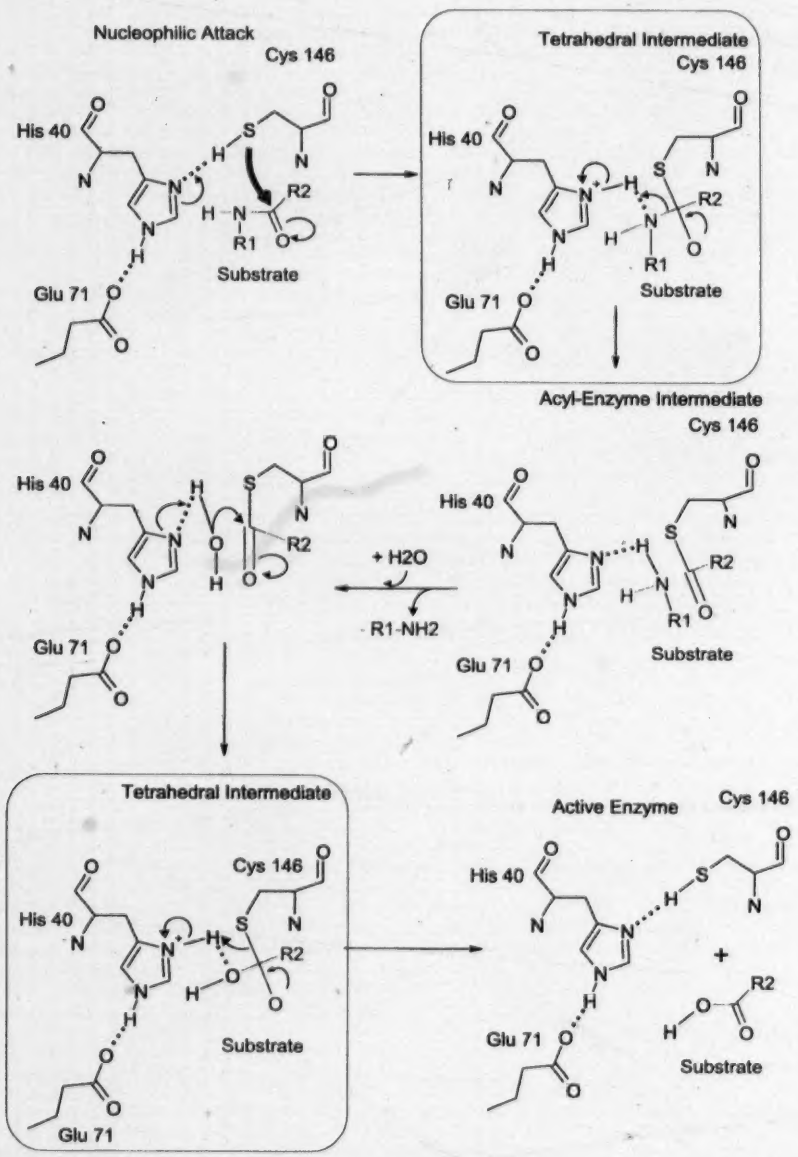

Figure 1.3: **Cysteine Protease Catalytic Mechanism.** The catalytic mechanism for cysteine proteases involve a Cys-His-Glu triad. Current opinion favors a tetrahedral intermediate that gives rise to an acyl-enzyme intermediate. Once the N-terminal peptide is replaced by an available water molecule, the enzyme is returned to its active state, while the cleaved peptide leaves with its new intact C-terminal carboxylic acid group.

sequence of six amino acids (TLFQ/GP). The HRV-3C proteases prefer a glutamine at the P_1 position, a glycine in the P_1' position and exhibit preference for a small, hydrophobic residue in the P_4 position [38]. Interestingly, the inclusion of proline in the P_2' position greatly increases substrate recognition [36, 37] and maintenance of residues to the P_5 position appears to increase cleavage efficiency.

This information has been utilized to direct the development of various irreversible peptidyl 3C protease inhibitors including ruprintrivir (formerly AG7088), which did entered clinical trials [39]. This inhibitor used an ethylpropenoate Michael acceptor group that undergoes nucleophilic attack. However, unlike an amide bond, its double bond undergoes rearrangement to form a stable adduct. This is graphically represented in Figure 1.4.

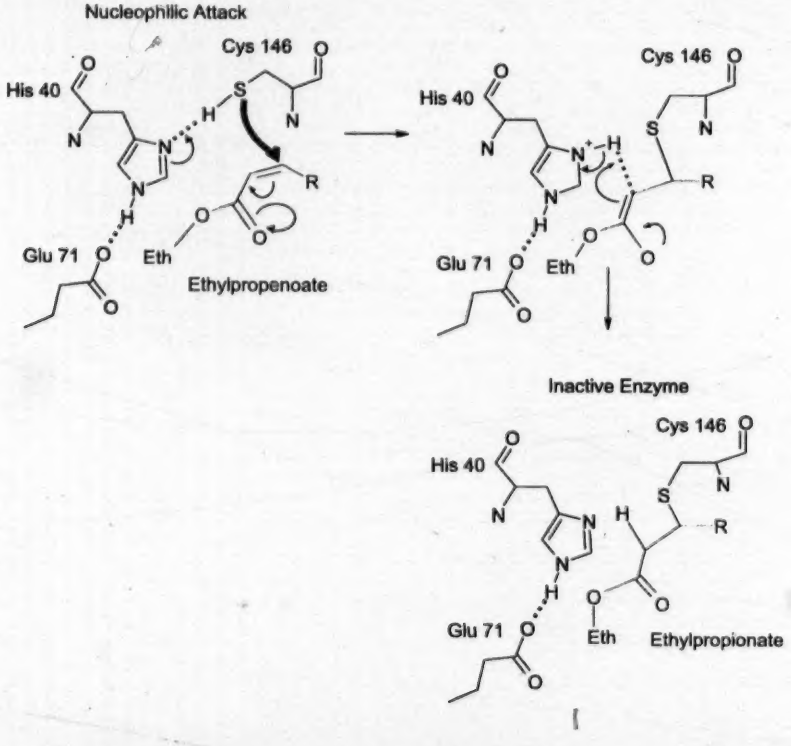

Figure 1.4: **3C Protease Inhibition with an Ethylpropenoate Michael Acceptor.** The catalytic mechanism for cysteine proteases is hindered with the covalent attachment of an ethylpropenoate group. A stable propionate group is formed following the covalent modification that irreversibly inactivates the enzyme.

In addition to this irreversible inhibitor, several peptide-mimic inhibitors that utilize re-

versible, competitive functional groups such as S-nitrosothiols [40] and alpha-ketoamides [41] have also been investigated. Interestingly, the structure-activity relationship studies done to date have shown some unexpected serotype and subgenus diversity in binding affinity and efficacy. Specifically, investigations into planar, hydrophobic residues in the substrate S_2 binding pocket presented different inactivation rates for the 3C proteases for serotypes 1A, 2, 10, 14, 16 and 89 [35]. These results were further investigated during irreversible-inhibitor optimization studies [42], which showed that hydrophobic moieties in the P_2 position were more favorable for serotype 14 proteases compared to serotypes 1A and 10. In light of this information, a better understanding of the proteolytic pharmacophore might aid in the development of a more universal lead pharmaceutical candidate.

To examine these sub-genus diversities further, the structure of one of the more widely studied variants of the rhinovirus, serotype 14, was chosen for investigation. However, it was necessary to first determine the three dimensional structure of the apo or free enzyme. Following this, a pharmacophore analysis of the proteolytic site could be conducted. Two commonly used techniques are currently available for high resolution macromolecular structure determination, X-ray crystallography and nuclear magnetic resonance (NMR). Other HRV-3C protease structures had been solved *via* X-ray crystallography. Our lab chose to use an alternative method, NMR. By using this method, we would obtain the first solution structure of any 3C protease to date and would be able to investigate or identify possible structural artifacts such as crystal packing. In addition to the tertiary structure, the chemical shifts of the protein would be obtained, which provide the necessary template to facilitate NMR perturbation studies like inhibitor screening and address the study of many unanswered biological questions such as heterodimeration interfaces and regulatory mechanisms. Because this study used NMR as the primary structural biological tool, a brief description of NMR, its fundamental principles, its experimental methodology and how this data is utilized in macromolecular structure calculation is warranted.

1.2 NMR Theoretical Principles

The nuclei of certain atomic isotopes can interact with a magnetic field and produce magnetic moments (μ). This property is dependant on an intrinsic property, the spin

angular momentum or I, which is quantized. Quantum mechanics dictates that each subatomic particle has a spin value of $\frac{1}{2}$. The combination of these spins from the subatomic constituents defines the overall spin property of an atom, whereby isotopes with an even number of neutrons and protons (ie. ^{2}H, ^{12}C) have spin angular momentums I of 0 and are not NMR active. Isotopes with an even to odd ratio of neutrons and protons (ie. ^{1}H, ^{13}C and ^{15}N) have spin angular momentums (I) of $\frac{1}{2}$ and isotopes with an odd number of neutrons and protons (*ie.* ^{14}N) have spin angular momentums (I) of 1. Each of these isotopes having non-zero spins are visible under NMR experimentation, however, the latter have complicated spin states and are not ideal for NMR observation. In particular, spins $> \frac{1}{2}$ have a non-spherical charge distribution that produces a quadrupole moment, which affects the isotope's relaxation time and broadens its NMR signal. Fortunately, elements with $I=\frac{1}{2}$ include the four most abundant organic atoms (H, C, N and O). The ^{1}H isotope of hydrogen has a natural abundance of >99%. The ^{13}C and ^{15}N isotopes of carbon and nitrogen, however, occur naturally with abundances of only 1.11% and 0.4% respectively. Observation of their NMR signal generally requires isotopic enrichment.

The isotopes with $I=\frac{1}{2}$ undergo a rotational motion or precession when an external magnetic field is applied. This rate of precession is known as the Larmor frequency (ω_0) and is proportional to the strength of the applied magnetic field (B_0) and an intrinsic property of the nucleus known as the gyromagnetic ratio (γ).

$$\omega_0 = -\gamma B_0 \tag{1.1}$$

Each nucleus has $2I+1$ possible spin orientation and $2I+1$ possible energy levels (E) known as Zeeman levels:

$$E = 2I + 1 \tag{1.2}$$

Nuclei such as ^{1}H, ^{13}C and ^{15}N have spin numbers of $\frac{1}{2}$ and can therefore adopt 2 possible energy levels defined by:

$$\Delta E = \gamma h B_0 / 2\pi \tag{1.3}$$

where h is Plank's constant (6.6363 x 10^{-34} J$*$s) and γ is the nuclear gyromagnetic ratio (1/s$*$T). Hence, the energy difference between the two energy levels is directly proportional

to the strength of the applied magnetic field (B_0) and the gyromagnetic ratio (γ). Different isotopes have different gyromagnetic ratios (γ) and therefore, have different sensitivities to an externally applied magnetic field. For example, ^1H, ^{13}C and ^{15}N have gyromagnetic ratios of 2.6752 x 10^8 s^{-1} T^{-1}, 6.728 x 10^7 s^{-1} T^{-1} and -2.712 x 10^7 s^{-1} T^{-1} respectively, making $^{13}C \sim \frac{1}{4}$ as sensitive and $^{15}N \sim \frac{1}{10}$ as sensitive as ^1H ($\propto \gamma^{\frac{5}{2}}$) when placed in the same magnetic field.

These nuclei will absorb a discrete amount of energy at their Larmor frequency when placed in the magnetic field. This energy will be divided between the $2I+1$ energy levels. The population difference between these energy levels is defined by the Boltzmann equation:

$$\frac{N_\beta}{N_\alpha} = exp^{(-\Delta E/kT)} \tag{1.4}$$

where N_β and N_α are the number of nuclei in the upper and lower energy states respectively, k is Boltzmann's constant and T is the temperature in Kelvin.

Collectively the nuclei create a net magnetization (M_0) due to the unbalanced population, which aligns with and precesses about the applied magnetic field (B_0) at a rate known as the Larmor frequency. A nucleus can absorb or emit energy when it is excited or when it resonates at its Larmor frequency. This energy is supplied in the form of electromagnetic radiation (a radio frequency pulse) that is applied perpendicular to the external B_0 field. This "pulse" generates a secondary magnetic field (B_1), which redirects the bulk magnetic momentum (M_0). By varying the phase, power and time of the electromagnetic "pulses", the bulk magnetic angular momentum can be directed in a plane transverse to the detector. This RF excitation gives rise to the NMR signal in the form of a decaying sinusoidal wave or free induction decay (FID). The intensity of this signal diminishes as the nuclei relax back to their ground state and realign with the magnetic field (B_0). An illustration of a single pulse NMR experiment and the simultaneous detection of all protons following a single broadband EM pulse is shown in Figure 1.5.

The detection of the time domain (ie. FID) data for all the excited nuclei can be converted to a frequency domain spectrum through Fourier transformation:

$$S(\omega) = \int_0^\infty s(t)\, e^{-i\omega t}\, dt \tag{1.5}$$

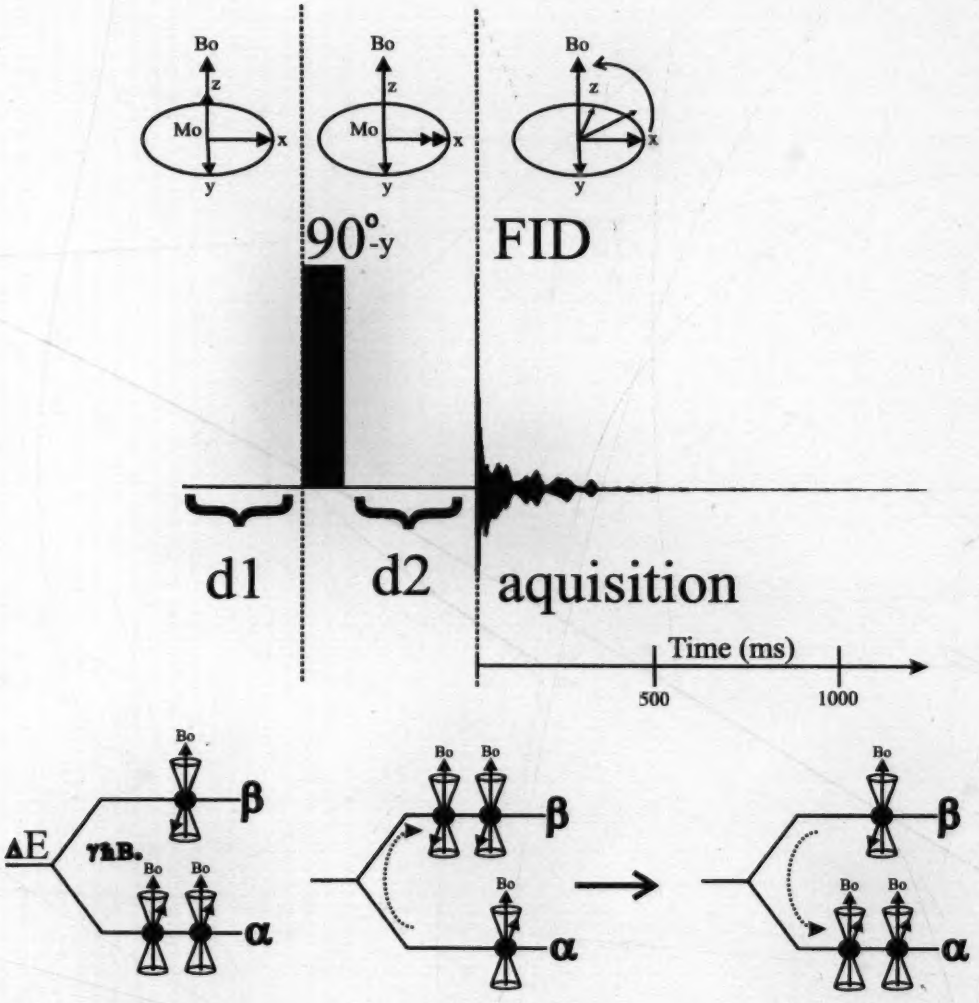

Figure 1.5: **1D NMR Experiment**. The free induction decay or FID of the ^{1}H atoms in the HRV14-3C protease following a single 1D pulse experiment. The experiment is divided into three phases. In the initial phase, 'd1', occurs before a RF pulse is applied. The protons are aligned with or against the applied magnetic field (B_0). Upon excitation with RF radiation at the NMR responsive atom's resonance frequency, the bulk magnetic moment (M_0) is aligned in a transverse plane relative to the NMR receiver coil ('d2' phase). As the small population excess of higher energy nuclei relax and realign back along the B_0 field, their precessional frequencies about B_0 induce decaying sinusoidal waves in the NMR coil during the acquisition phase. This analogue signal or FID from the multiple frequencies is digitized and Fourier transformed to present the customary NMR spectrum.

where $S(\omega)$ is the converted frequency from the time domain signal $s(t)$ and has quantity with both amplitude and phase properties.

A vast amount of data about a biological macromolecule can be obtained from an increasingly large number of NMR experiments. Many of these experiments are used to derive what are known as chemical shifts. Additional types of NMR data such as NOEs and coupling constants are dependant on structure. These kinds of data are usefull for macromolecular structure calculations in that they provide restraint information. These NMR derived data will be discussed briefly.

1.2.1 Chemical Shifts

An NMR spectrum consists of several distinct features. The most prominent of these consist of peaks or clusters of peaks located at distinct positions or characteristic frequencies. These characteristic frequencies are called chemical shifts, δ, which result because of the different chemical environments experienced by each nucleus. The electron density surrounding each nucleus effectively "shields" the nucleus from the applied magnetic field, B_0, *via* their own magnetic fields. This leads to each nucleus experiencing slightly different magnetic field strengths, B_1, defined by:

$$B_1 = B_0(1 - \sigma) \tag{1.6}$$

where σ is the "shielding" constant. This constant is comprised of a number of terms, which includes but is not limited to electron density, paramagnetic, diamagnetic, neighbor and ring current effects. These combined influences alter each nucleus electronic environment, thus differentiating their perceived B_O and their precessional or Larmor freqency. These differences are what give rise to the different peaks that comprise a complex spectrum (Figure 1.6), rather than a single freqency or single peak for each NMR sensitive nucleus.

The differences in these frequencies is relatively small. Therefore, it is customary to report each nucleus' precessional frequency as a chemical shift, σ. The chemical shift if measured in terms of parts per million (ppm) relative to a chemical shift standard such as DSS. This value is thus independent of the B_0 and is comparable with values collected on different spectrometers with different magnetic field strengths. By convention the ppm

values are calculated with the following formula:

$$\delta_{(ppm)} = \frac{\upsilon_{observed} - \upsilon_{reference}}{\upsilon_{spectrometer}} \; x \; 10^6 \; _{(Hz)} \tag{1.7}$$

where v are the measured rotational frequencies in Hertz. An incredible amount of information is available from chemical shift analysis. The very fact that chemical shifts are present in different areas of the NMR spectrum can identify the various types of nuclei (^1H_N, $^1H_\alpha$, $^{13}C_\alpha$, etc.) present. Furthermore, because they are sensitive to their local environment, slight changes in their positions within their usual or characteristic spectral regions can indicate their local secondary structure [43] and provide a wealth of knowledge regarding cysteine redox states [44], histidine tautomeric states [45] and hydrogen bond lengths [46].

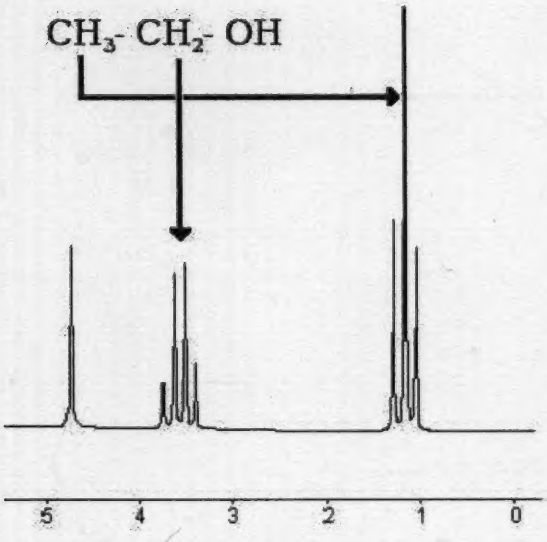

Figure 1.6: **1D ^{1}H NMR Spectrum of Ethanol.** The 1D NMR spectrum shows the different peaks observed for the different protons covalently bound to the methyl and methylene carbons on ethanol. The different electronic environments experienced by the protons resulting from their proximity to the de-shielding hydroxyl group gives rise to the different chemical shifts, shown here to be ~1.1 ppm and ~3.5 ppm for the methyl and methylene protons respectively.

1.2.2 Coupling Constants

Another important NMR phenomena are coupling constants or J, which are manifested as split peaks or fine structure "multiplets" (Figure 1.7). These manifestations result in splitting of a given peak for an NMR active nucleus when it neighbors non-equivalent nuclei (generally $\leq$ 3 bond lengths). The low and high energy states (α and β) that are adopted for the neighboring spin $\frac{1}{2}$ nuclei influence the magnetic field of the observed nuclei such that slight variations of its Larmor frequency result.

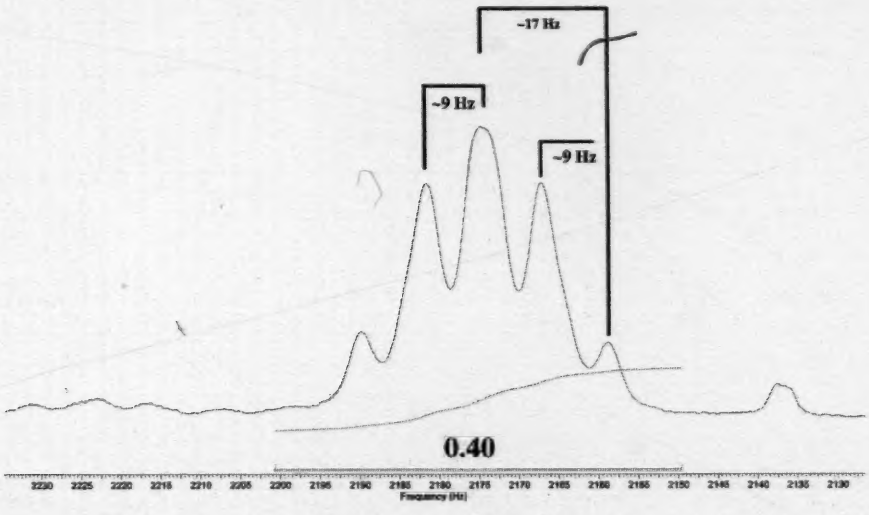

Figure 1.7: ***J*-Coupling Constants.** The fine splitting observed in a coupled 1D ^{1}H NMR spectrum is depicted in this peak.

These differences in the Larmor freqency result is different chemical shifts or "peak splitting". These constants are magnetic field strength independent and reflect the rate at which nuclear magnetization is transfered between NMR sensitive nuclei, such as ^{1}H and ^{15}N, through their covalent bonds. The existence of *J*-couplings allowed for the development of a number of NMR experiments that can indirectly detect covalently linked nuclei. Examples of these experiments include the ^{15}N-HSQC, ^{1}H-^{1}H-TOCSY [47], HNCO [48] and HNCACB [49] experiments. These 2D and 3D experiments transfer magnetization between adjacent, "coupled" nuclei. It is through experiments such as these that respective atoms within the residue assigned specific chemical shifts and 'spin systems' for individual residues within a protein sequence can be linked. The individual *J*-coupling contants can be measured directly from experiments like the 2D ^{1}H-^{1}H-COSY [50] or indirectly from experiments like the 3D HNHA [51]. Equally important is the fact that these couplings are contingent on the orientation between the bonds that covalently link the atoms. This relationship is useful in defining the geometric orientation between the atoms and can be used to define torsion angular restraints during molecular structure calculation. An example of

this information derivation is given by the Karplus equation [52]:

$$^3J_{ab} = A * cos^2(\theta - \phi) + B * cos(\theta - \phi) + C \tag{1.8}$$

where ϕ is the interatomic three-bond dihedral angle, θ is the coupling dependant phase and ^3J is the three-bond coupling constant between atoms a and b. A, B and C are empirically derived constants optimized for each kind of coupling. For example, ϕ is -60° and constants of 6.98, 1.38 and 1.72 are used for calculating the phi dihedral angle restraints in proteins from $^3J_{HNH\alpha}$ couplings [51]. This is represented graphically in (Figure 1.8).

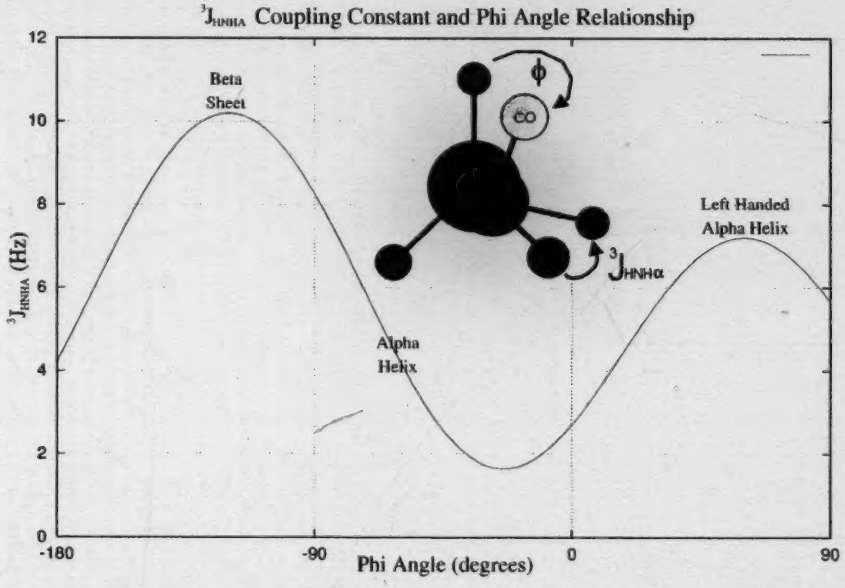

Figure 1.8: $^3J_{HNH\alpha}$ **Karplus Equation Relationship.** The relationship between calculated ^3J HNHα coupling constants and the the bond dihedral angle that connects the atoms is represented by the graph. Inserted in the graph is a diagram showing the ϕ dihedral angle formed between the bonds. Certain types of secondary structure have distinct ϕ angles (β-strands: $\sim$ -120°, α helices: $\sim$ -60°) and therefore, give rise to distinct ^3J-coupling constants (>8 Hz and $\leq$ 4 Hz respectively).

1.2.3 Nuclear Overhauser Effect

Besides the through bond-transfer of magnetization (J-coupling) that was just described, magnetization can also be transfered through space by saturating the higher energy state nuclei. These saturated nuclei can thus transfer magnetization to a neighboring nucleus ($\leq$5 Å) resulting in a net enhancement of the NMR signal among the neighboring nuclei. This correlation, termed the Nuclear Overhauser Effect or NOE is inversely related to the interacting protons (i and j) inter-atomic distance by a factor of d_{ij}^{-6}. This spatially influenced signal enhancement fades quickly and is used to derive inter-proton distance restraints or NOEs from experiments like the 3D ^{15}N -edited NOESY [53] and 3D ^{13}C-edited NOESY [54]. This information forms the cornerstone of macromolecular structure determination by NMR. Long range ($i|j > 4$ residues apart) NOEs aid in determining the three dimensional 'fold' of a given macromolecule, which can be derived from methyl to methyl proton contacts and characteristic weak, medium and strong NOE patterns between spatially close hydrogen atoms. For example, anti-parallel β-strands exhibit characteristic cross-strand NOEs that include strong and medium intensity NOE signals from the corresponding $d_{\alpha\alpha}$ (~2.2 Å) and d_{NN} (~3.6 Å) distances between the interacting protons [55]. This is illustrated in Figure 1.9.

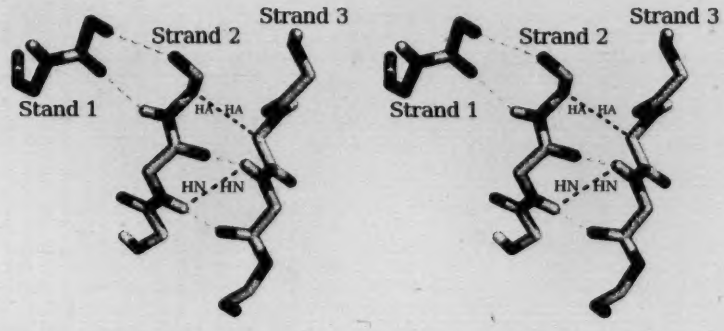

Figure 1.9: **β-Strand Secondary Structure**. The positional relationship of inter-residue protons residing on adjacent anti-parallel β-strands is shown in this wall-eyed stereo image. The network of hydrogen bonds between adjacent strands are shown by broken yellow lines. The typical cross-strand H_N-H_N and H_α-H_α NOEs observed in NMR spectra that result from this secondary structure are represented by broken blue lines. H_α hydrogens are not shown for clarity.

Table 1.1: **T_1, T_2 and τ_c vs, Protein Molecular Weight†.**

τ_c (ns) 35 °C	Mwt (kDa)	^{15}N T_1 (ms)	^{15}N T_2 (ms)	^{13}C T_1 (ms)	^{13}C T_2 (ms)
5	10	410	150	410	40
7	12	500	110	550	30
9	17	600	90	700	25

†Adapted from *Protein NMR Techniques*[56]

1.2.4 Physical Parameter Considerations

Another characteristic of NMR spectra are the shapes or widths of the peaks (Figure 1.10). In particular, the linewidth or "*sharpness*" of NMR peaks depend on the size of the molecule and the rotational correlation time (τ_c) of the molecule being studied. Two relaxation times (T_1 and T_2) are related to this intrinsic parameter. The relationship between T_2, T_1, and τ_c for proteins of varying sizes is given in Table 1.1.

A nucleus's spin-lattice (T_1) relaxation time correlates proportionally with the size of the molecule and is the time needed for a nucleus to return to equilibrium following excitation by a RF 'pulse'. Therefore, a protein's T_1 time determines the recycle delay between collecting consecutive FIDs and therefore, the amount of data collected on a sample in a given time period. This is important because the more FID data that is averaged together improves the signal to noise of any given NMR experiment.

The τ_c is related inversely to the spin-spin relaxation time (T_2), therefore, higher molecular weight species exhibit reduced spectral quality due to increased spectral line widths (Figure 1.10). Increasingly larger line-widths result in spectral overlap and make measurement of spectral data like coupling constants and NOEs difficult. Figure 1.10 indicates that a key factor for obtaining good NMR data is related to the size of the protein being studied. Although the molecular weight of a protein is intrinsic, it can be artificially inflated by self association or inter-protein interactions, which significantly increase τ_c and reduce T_2. These detrimental consequences, however, can sometimes be controlled through sample purification conditioning. For these reasons, determining sample conditions that maximize the T_2 time (ie. maintain a monomeric state) are important for obtaining high quality NMR data.

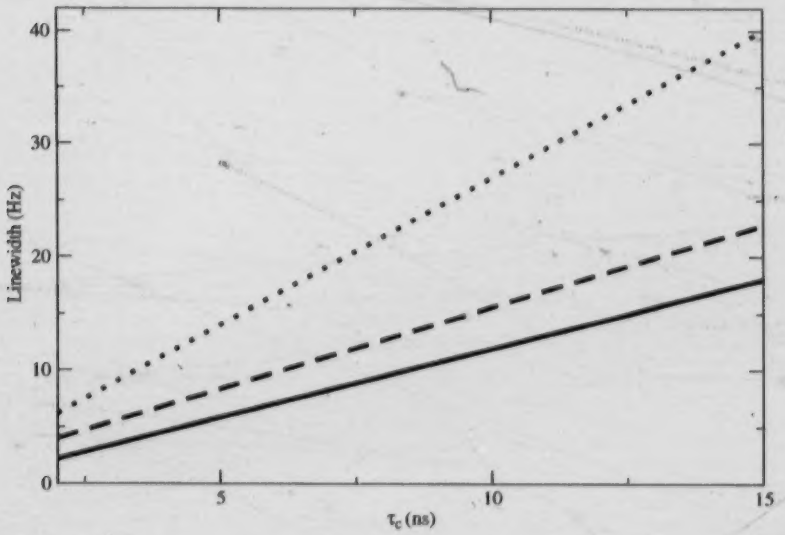

Figure 1.10: **Protein τ_c vs. Spectral Linewidths.** Adapted from *Protein NMR Spectroscopy: Principles and Practice* [57]. The relationship between increasing τ_c and and increasing linewidths is linearly proportional. Linewidths for protons in an unlabelled sample (—), for protons bound to ^{13}C ($\cdots$) and ^{15}N ( ).

1.3 Macromolecular Structure Determination

Before many structural biology questions can be addressed, it is often necessary to have a macromolecular structure of a protein, DNA or carbohydrate in hand. There are many questions that can be answered from low resolution structures or global folds . Examples include functional classification and macromolecular structure interface identification, however, other questions like the analysis of a protein's pharmacophore or detailed structural interactions with a ligand require more detailed, high resolution structures. To obtain such data a series of experimental steps are generally followed, which are itemized:

1. Sample Preparation
2. Data Collection
3. Sequential Assignment and Validation
4. Structure Calculation and Validation

Each of these steps is equally important, however, first and foremost is the procurement of a concentrated, protein sample that is also stable at conditions conducive to collecting NMR spectra. Weakly concentrated samples, with or without contaminants and poor stability can end the NMR study of a protein prematurely. It is absolutely critical that the best available samples be used for the collection of NMR data. This will afford the best possible data, will aid the spectral analysis and expedite the project.

1.3.1 NMR Sample Preparation

High sample purity is not only necessary to reduce any signals that might come from contaminant sources but also minimize possible protein-protein interaction. Furthermore, NMR sample stability must exceed the time required to collect a standard three dimensional heteronuclear NMR experiment, which is on the order of 72 to 96 hours (3 to 4 days). Factors that seed aggregation or precipitation need to be addressed as these processes reduce the amount of material available for NMR detection (< mM).

Interactions with contaminant sources can be detrimental if the contaminant is a protease. Fragments resulting from proteolytic processes can reduce the amount of material available for detection, seed precipitation and produce additional signals that increase spectral complexity that can further complicate the assignment process. Because the HRV14-3C protein is proteolytic, auto-lysis requires consideration during protein purification and sample preparation.

Experimental conditions such as concentration, pH and temperature, although dictated by a given proteins stability, can affect the NMR data quality and can limit the viability of any NMR study. Obviously, increasing the protein concentration directly increases the number of nuclei available for detection. However, increasing the concentration can also increase the likelihood of protein self-association. Therefore, the maximum possible concentration a protein can tolerate without seeding precipitation needs to be determined prior

to starting any heteronuclear NMR experiment.

A number of heteronuclear NMR experiments involve detecting nuclear frequencies indirectly by transferring NMR induced magnetization to the solvent exchangeable backbone amide proton. Sample conditions that minimize the exchange process, like acidic pH, are important considerations when optimizing sample conditions. Temperature is another important parameter that requires balance and compromise when conducting NMR experiments. Sample temperature affects the quality of NMR data in a number of ways. Although decreasing the sample temperature increases NMR sensitivity (Equation 1.4), increasing the temperature decreases the protein rotational correlation time (τ_c), thereby improving spectral linewidths (Table 1.1). Furthermore, increasing the sample temperature can increase the number of spin systems undergoing fast exchange and thus increases the number of visible spin systems by reducing the number of nuclei undergoing intermediate exchange. Therefore, deriving sample parameters that maximize the stability of protein samples at above room temperature conditions is often beneficial. These parameters were taken into consideration when developing the expression and purification protocols for the HRV14-3C protease and optimizing its solution state conditions.

1.3.2 Data Collection and Sequential Assignment

Structure determination of smaller proteins (<10 kDa) can often be completed with the use of 1D and 2D proton spectroscopy. This work can be readily conducted on the naturally abundant and highly sensitive ^{1}H nuclei. However larger proteins like HRV14-3C, suffer from severely overlapped spectra and require enrichment of ^{15}N and ^{13}C isotopes. Enrichment with these isotopes improves spectral resolution by allowing separation of the spin systems between the various nuclear 'dimensions'. This concept is represented in Figure 1.11.

Usually, an ensemble of heteronuclear NMR experiments are collected and analyzed simultaneously to piece together the intra and inter-residue connectivity within a protein macromolecule. A variety of complimentary 2D and 3D NMR experiments like ^{15}N-HSQC, HNCA [58], HNCACB [49] and C(CO)NNH [59] provide both intra (i) and inter-residue ($i-1$) chemical shift information (*ie.* δ of i $^{1}H_N$ to i and $i-1$ $^{13}C_\alpha$ and $^{13}C_\beta$ nuclei). This "overlap" is used to connect each residue's distinct spin system and sequentially link

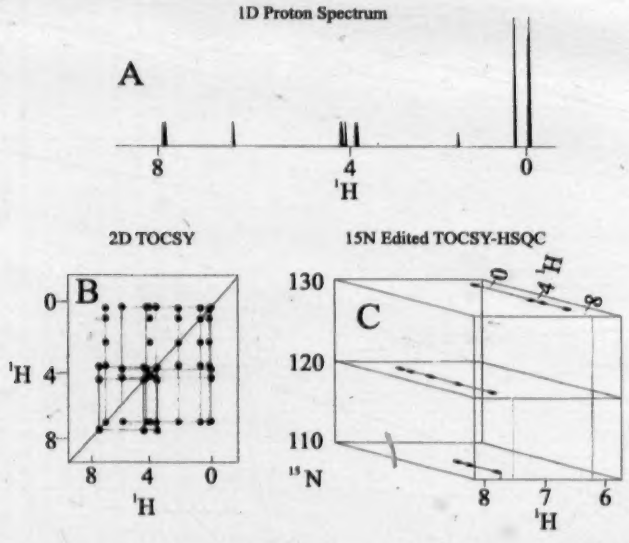

Figure 1.11: **Multiple Dimension NMR Experiments.** Illustrated are representations of one, two and three dimensional NMR experiments for as tripeptide (Gly-Val-Thr). The glycine NMR peaks are colored blue, the valine peaks are colored green and the threonine peaks are colored red. A number of NMR peaks are overlapped in the 1D experiment (**A**), including the Hα's for all the residues and the Hγ's for the valine and threonine residues. These peaks become somewhat resolved in the 2D-^{1}H-TOCSY experiment (**B**) when correlated along the H_N region. However, the H_α and H_γ regions still remains overlapped. When these spin systems are further resolved along the ^{15}N-dimension in the 3D ^{15}N-edited TOCSY-HSQC (**C**), complete resolution of the three spin systems becomes clear. All axes are in ppm.

them. These distinct 'blocks' of spin systems can then be assigned to regions within the protein sequence, thereby completing the protein's chemical shift assignments. In addition to these inter-residue connectivity experiments, NMR spectra like the HCCH-TOCSY [60] and CCH-TOCSY [61, 62] are useful in providing and completing intra-residue side-chain information.

Collectively, this process is called sequential assignment and must be completed prior to any further NMR analysis. Generally, this process is straight-forward, however, it can sometimes present problems if regions of the protein being studied undergo intermediate chemical exchange. An example of the assignment process using a variety of NMR experiments is shown in Figure 1.12.

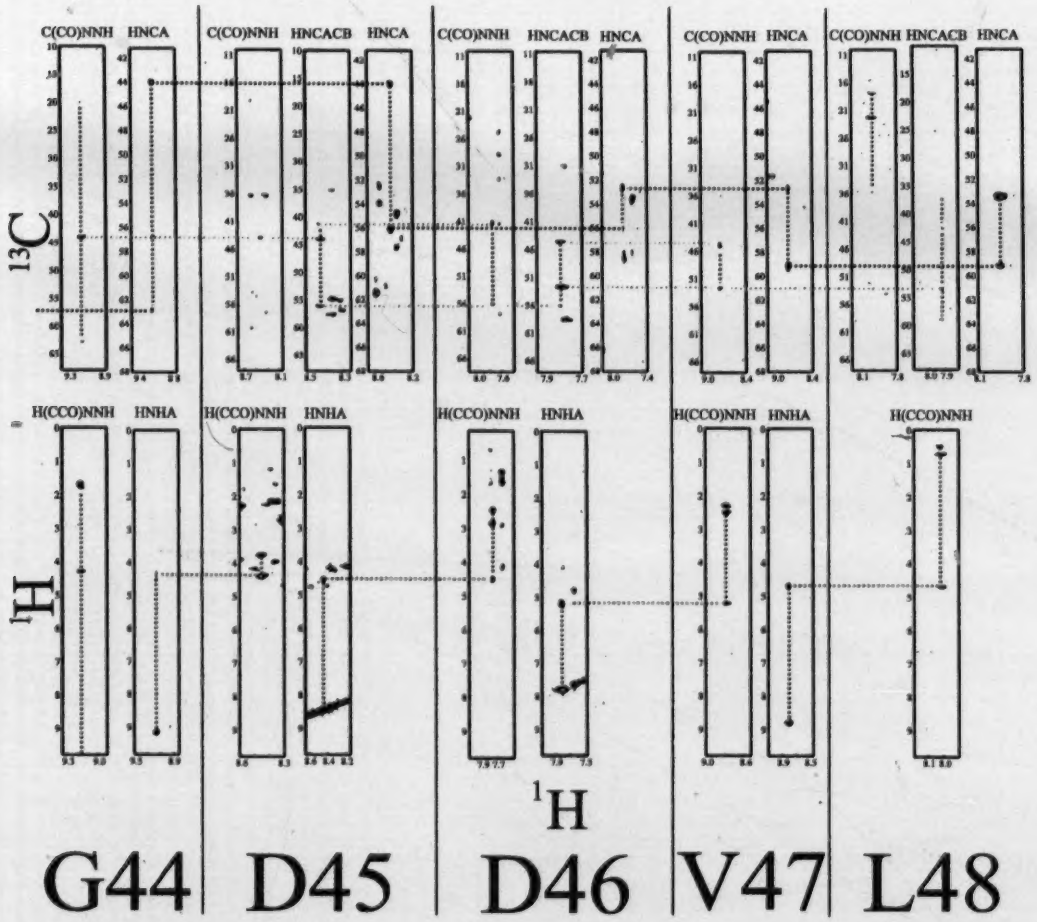

Figure 1.12: **Sequential Connectivity**. An example of how a residue's '*spin system*' is completed and linked sequentially to adjacent residues within a protein sequence by using combinations of NMR experiments. The x-axis for all spectra is the amide proton dimension. All axes are in ppm. Broken blue lines join intra-residue NMR peaks. Broken black lines connect $^{13}C_\alpha$ i and $i-1$ peaks within the HNCA experiment between sequentially adjacent residues. Broken green lines connect $^{13}C_\alpha$ and $^{13}C_\beta$ i and $i-1$ peaks between sequentially adjacent residues in the C(CO)NNH and HNCACB experiments. Broken red lines connect i and $i-1$ $^{1}H_\alpha$ peaks in the HNHA and H(CCO)NNH experiments between sequentially adjacent residues. Strips in the top row show examples from HNCA, HNCACB and C(CO)NNH spectra, which are used to assign carbon chemical shift data to a given residue. Some of these experiments (HNCA and HNCACB) provide intraresidue (i) and preceding residue ($i-1$) chemical shift information. Other experiments (*ie.* C(CO)NNH) give only preceding ($i-1$) information. Overlaying this data allows for the chemical shifts to be assigned to a particular residue within the protein sequence. The strips shown in the bottom row are from HNHA and H(CCO)NNH experiments. These experiments are used to complete the proton chemical shift assignments for a given residue and can also be used to aid in the sequential assignment.

1.3.3 Structure Calculation

Once the chemical shifts of a protein are assigned, structurally dependent data (NOEs) can be derived from NOESY spectra. These spectra typically resemble TOCSY spectra due to the close proximity of intra-residue nuclei. However, closer inspection reveals a number of additional signals that arise from distinct, non-covalently bound protons (Figure 1.13).

These NMR signals are important for deriving inter-proton restraints between nuclei that may be quite distal sequentially but quite close spatially. This is represented in Figure 1.14. By gathering these data and quantizing them, a set of inter-proton distance restraints or NOEs can be obtained and entered into programs like XPLOR [63] and CYANA [64]. These programs use distance geometry and simulated annealing protocols to 'fold' the molecule into a three-dimensional structure that satisfies the input distance data (*ie.* no violations). Generally an ensemble of structures is generated and the structures presenting the lowest energy scores with no or minimal violations are selected to represent the experimental structure. Energy terms are usually calculated empirically by comparing geometric parameters such as bond lengths, bond angles and van der Waals contacts to 'idealized' parameters. Although NOEs form the basis of these calculations, other data like dihedral angle restraints (derived from ^3J-coupling constants) and hydrogen bond restraints can be used to improve convergence after the preliminary structures are determined. Chemical shifts can also be used to derive torsion angles. Furthermore, refinement in explicit solvent can improve geometric parameters and overall structure quality [65].

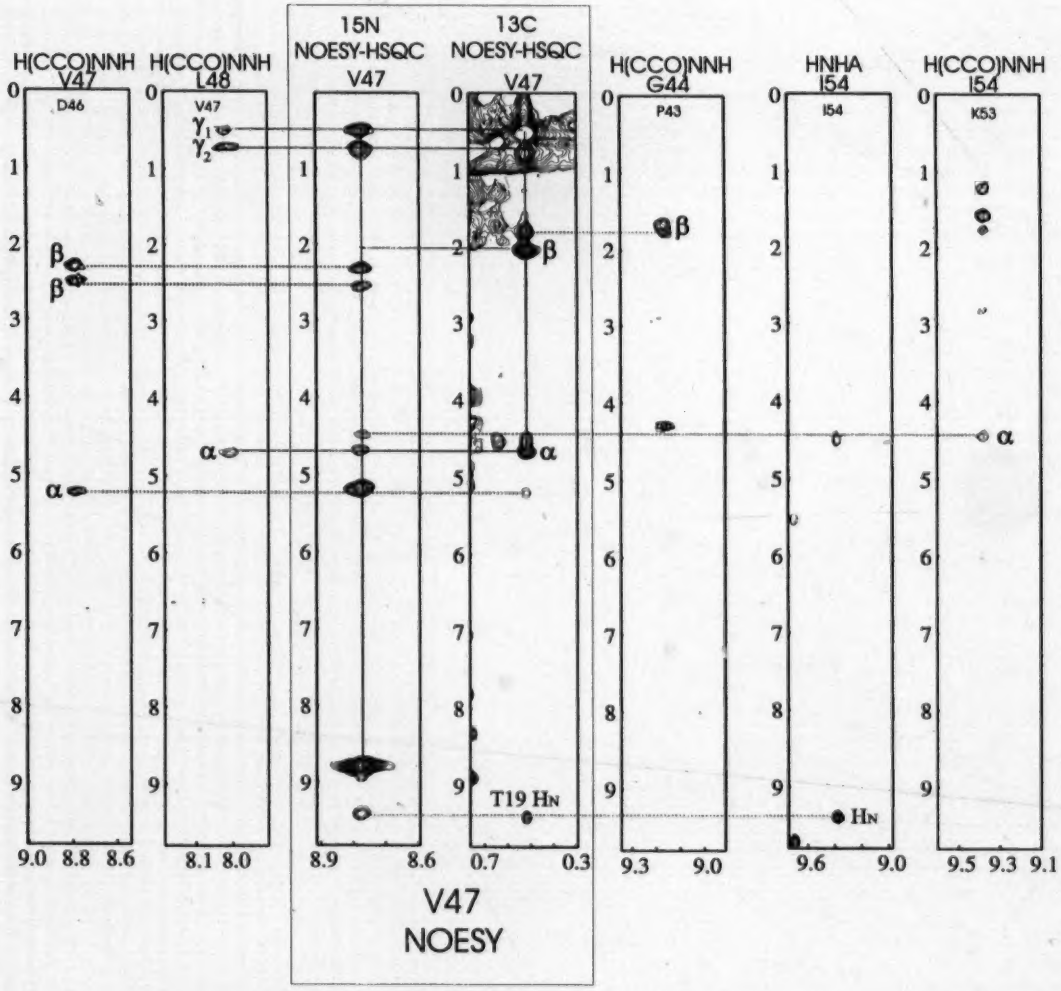

Figure 1.13: **NOE Data Derivation**. All axes shown are for the proton dimensions and are in ppm. The various strip plots show data from a variety of different spectra for a number of spatially close residues within the HRV14-3C protease. The labels inserted in each strip plot represents the specific residue that each strip plot provides chemical shift information for. The two boxed strip plots show NOE data for Val^{47}. Val^{47}'s "*spin system*" is connected by orange lines. The extra signals within the NOESY plots arise from inter-residue NOEs. Broken lines connect these NOE peaks to "*spin system*" peaks derived from the *J*-coupled experiments. Blue lines connect NOEs to Val^{47}'s amide proton (obtained from the ^{15}N-edited NOESY-HSQC experiment). Red lines connect NOEs to one of Val^{47}'s $^{1}H_{\gamma}$ protons (obtained from the ^{13}C-edited NOESY-HSQC experiment). These NOEs correlate to both sequential (*i=1*) and distal (*i>1*) NOEs and are mapped onto the structure shown in Figure 1.14.

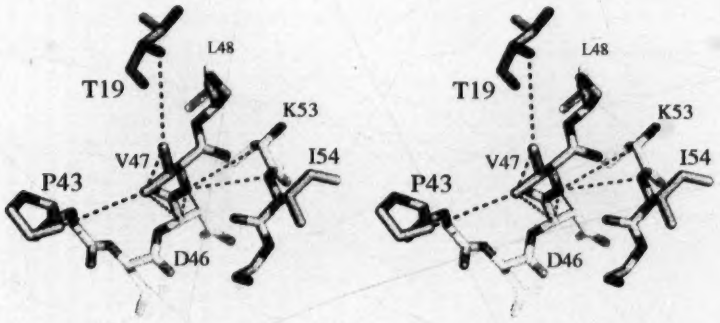

Figure 1.14: **NOEs for Val^{47}**. Wall-eyed stereo view of spatially adjacent residues to Val^{47}. Blue broken lines show NOEs obtained from the ^{15}N-edited NOESY-HSQC spectrum and red broken lines show NOEs obtained from the ^{13}C-edited NOESY-HSQC spectrum. The strip plots from these spectra used to assign these NOEs are shown in Figure 1.13. Protons have been removed from this structure for clarity. The broken lines depicting NOEs are drawn from the heavy atoms, C (red) and N (blue), bonded to the protons. The spin diffusion artifact observed for the $^{1}H_{\gamma 1}$ protons of Val^{47} (through the $^{1}H_{\gamma 2}$ protons) to the $^{1}H_{N}$ of Thr^{19} is shown in green.

1.4 Thesis Outline

To gain a better undertanding of the sub-genus diversity observed for the HRV-3C proteases, we used NMR as an experimental method for the structure determination of the HRV14-3C protease both with and without a covalently bound peptide inhibitor. Although the structure of the apo form was first described 12 years ago, this structural data has never been released. Another key motivating factor to complete the HRV14-3C protease structure lies with the fact that it belongs to the rhinovirus sub-genus B group and its structure would allow detailed comparison with the sub-genus A, HRV2-3C protease, whose X-ray structural data have been released.

In addition to the introduction, this dissertation consists of five additional chapters that describe the preparation of the protein, the NMR sample conditions, the synthesis of the peptide inhibitor, the solution structure of the inhibited form of the enzyme and the solution structure of the apo form of the enzyme. More specifically, Chapter 2 outlines the methods used to express the protein in *E. coli* cells and subsequently purify it, in which care was taken to minimize possible deamidation products that had been identified in previ-

ous studies. Following this, solution conditions were examined that would be conducive to collecting NMR spectra (ie. high concentration, low salt, etc.) that optimized sample stability. Chapter 3 discusses the synthesis of an inhibitor for the HRV14-3C protease. Initial sequential assignment work on the protease was difficult due to weak NMR signals and poor sample stability (~1 week). This redirected the research toward improving stability and NMR spectra quality. The improved results were far greater than expected and ultimately led to the completion of the NMR structure of the inhibited enzyme. This work is presented in chapter 4, which outlines all the NMR experiments used to obtain the backbone, side-chain and restraint data. The methods used to assign the protein chemical shifts are also presented. Likewise the structure calculation and quality assessment is outlined. Finally, this chapter analyzes the pharmacophore of the enzyme and compares it to the homologue HRV2-3C protease. New insights into upstream scissile bond interactions are also presented and discussed. Chapter 5 addresses the chemical shift assignment and solution structure generation of the apo HRV14-3C protease. New purification steps improved the sample stability and the now-known, inhibited structure provided a necessary template to complete the apo structure. Alternative methods were also used to calculate this structure. These methods are described along with a structural and dynamic assessment between the apo and bound forms of the enzyme. The thesis concludes with a brief discussion of possible future directions this project might take.

Bibliography

[1] Anonymous. The Common Cold. NIAID Fact Sheet [online]. `http://www.niaid.nih.gov/factsheets/cold.htm`.

[2] P. W. Heymann, T. A. Platts-Mills, and S. L. Johnston. Role of Viral Infections, Atopy and Antiviral Immunity in the Etiology of Wheezing Exacerbations Among Children and Young Adults. *Pediatr Infect Dis J*, 24:217–222, 2005.

[3] M. J. Mäkelä, T. Puhakka, O. Ruuskanen, M. Leinonen, P. Saikku, M. Kimpimäki, S. Blomqvist, T. Hyypiä, and P. Arstila. Viruses and Bacteria in the Etiology of the Common Cold. *J Clin Micro*, 36:539–542, 1998.

[4] C. Y. Hong, R. T. Lin, E. S. Tan, P. N. Chong, Y. S. Tan, Y. J. Lew, and L.H. Loo. Acute Respiratory Symptoms in Adults in General Practice. *Fam Prac*, 21:317–323, 2004.

[5] M. T. Coiras, J. C. Aguilar, M. L. Garcia, I. Casas, and P. Perez-Brena. Simultaneous Detection of FFourteen Respiratory Viruses in Clinical specimens by Two Multiplex Reverse Transcription Nested-PCR Assays. *J Med Virol*, 72:484–495, 2004.

[6] R. Ostroff, A. Ettinger, H. La, M. Rihanek, L. Zalman, J. Meador III, A. K. Patick, S. Worland, and B. Polisky. Rapid Multiperotype Detection of Human Rhinoviruses on Optically Coated Silicon Surfaces. 21:105–117, 2001.

[7] H. Dagher, H. donninger, P. Hutchinson, R. Ghildyal, and P. Bardin. Rhinovirus Detection: Comparison of Real-time and Conventional PCR. *J Virol Meth*, 117:113–121, 2004.

[8] M. G. van Kraaij, L. J. van Elden, A. M. van Loon, K. A. Hendriksen, L. Laterveer, A. W. Dekker, and M. Nijhuis. Frequent Detection of Respiratory viruses in Adult Recipients of stem Cell Transplants with the Use of Real-time Polymerase Chain Reaction, Compared with Viral Culture. *Clin Infect Dis*, 40:662–669, 2005.

[9] S. Bellau-Puljol, A. Vabret, L. Legrand, J. Dina, S. Gouarin, J. Peitjean-Lecherbonnier, B. Pozzetto, C. Givevra, and F. Freymuth. Development of Three Multiplex RT-PCR Assays for the Detection of 12 Respiratory RNA Viruses. *J Virol Meth*, 126:53–63, 2005.

[10] J.M. Gwaltney, P. B. Moskalski, and J. O. Hendley. Hand-to-hand Transmission of Rhinovirus Colds. *Annals of Intern Med*, 88:463–467, 1978.

[11] J.M. Jr. Gwaltney and J. O. Hendley. Transmission of Experimental Rhinovirus Infection by Contaminated surfaces. *Am J Epidem*, 116:828–833, 1982.

[12] A. Jadwiga and M.D. Wedzicha. Exacerbations*: Etiology and Pathophysiologic Mechanisms. *Chest*, 121:136S–141S, 2002.

[13] R. Jr. Lemanske, D. Jackson, R. Gangnon, M. Evans, Z. Li, P. Shult, C. Kirk, E. Reisdorf, K Roberg, and Anderson E. Rhinovirus Illnesses During Infancy Predict subsequent Childhood Wheezing. *J Allergy and Clin Immunol*, 116:571–577, 2005.

[14] T. Ohlmann, M. Rau, V. M. Pain, and S. J. Morley. The C-terminal Domain of Eukaryotic Protein Synthesis Initiation Factor (eIF) 4G is Sufficient to Support Cap-independent Translation in the Absence of eIF4E. *Embo J*, 15:1371–1382, 1996.

[15] R. Sharma, S. Raychaudhuri, and A. Dasgupta. Nuclear Entry of Poliovirus Protease-Polymerase Precursor 3CD: Implications for Host Cell Transcription Shut-off. *Virology*, 320:195–205, 2004.

[16] B. L. Dasgupta, A. *et al.* (eds. Semler and E.) Wimmer. *Molecular Biology of Picornaviruses*. ASM Press, Washington, DC, USA, 2002.

[17] R. Banerjee, M. K. Weidman, A. Echeverri, P. Kundu, and A. Dasgupta. Regulation of Poliovirus 3C Protease by the 2C Polypeptide. *J Virol*, 78:9243–9256, 2004.

[18] C Badorff, G. H. Lee, B. J. Lamphear, M. E. Martone, K. P. Campbell, R. E. Rhoads, and K. U. Knowlton. Enteroviral Protease 2A Cleaves Dystrophin: Evidence of Cytoskeletal Disruption in an Acquired Cardiomyopathy. *Nat Med*, 5:320–326, 1999.

[19] J. R. Doedens and K. Kirkegaard. Inhibition of Cellular Protein Secretion by Poliovirus Proteins 2B and 3A. *Embo J*, 14:894–907, 1995.

[20] S. B. Deitz, D. A. Dodd, S. Cooper, P. Parham, and K. Kirkegaard. MHC 1-dependent Antigen Presentation is Inhibitied by Poliovirus Protein 3A. *Proc Natl Acad Sci U S A*, 97:13790–13795, 2000.

[21] P. A. Walker, L. E-C. Leong, and A. G. Porter. Sequence and Structural Determinants of the Interaction between the 5'-Noncoding Region of Picornavirus RNA and Rhinovirus Protease 3C. *J Biol Chem*, 270:14510–14516, 1995.

[22] W. Xiang, A. Cuconati, A. V. Paul, X. Cao, and E. Wimmer. Molecular Dissection of the Multifunctional Poliovirus RNA-binding Protein 3AB. *RNA*, 1:892–904, 1995.

[23] J. S. Towner, T. V. Ho, and B. L. Semler. Determinants of Membrane Association for Poliovirus Protein 3AB. *J Biol Chem*, 271:26810–26818, 1996.

[24] A. V. Paul, X. Cao, K. S. Harris, J. Lama, and E. Wimmer. Studies with Poliovirus Polymerase $3D^{pol}$. Stimulation of poly(U) Synthesis *in vitro* by Purified Poliovirus Protein 3AB. *J Biol Chem*, 269:29173–29181, 1994.

[25] H. H. Roehl and B. L. Semler. Poliovirus Infection Enhances the Formation of Two Ribonucleoprotein Complexes at the 3′ End of Viral Negative-strand RNA. *J Virol*, 69:2954–2961, 1995.

[26] K. E. Gustin and P. Sarnow. Effects of Poliovirus Infection on Nucleocytoplasmic Trafficking and Nuclear Pore Complex Composition. *Embo J*, 20:240–249, 2001.

[27] A. L. Hughes. Phylogeny of the Picornaviridae and Differential Evolutionary Divergence of Picornavirus Proteins. *Infec Genet and Evol*, 4:143–152, 2004.

[28] W. Olszewska, M. Zambon, and P. J. M. Openshaw. Development of Vaccines Against Common Colds. *Brit Med Bul*, 62:99–111, 2002.

[29] F. G. Hayden, D. T. Herrington, T. L. Coats, K. Kim, E. C. Cooper, S. A. Villano, S. Liu, S. Hudson, D. C. Pevear, M. Collett, and M. McKinlay. Efficacy and Safety of Oral Pleconaril for Treatment of Colds Due to Picornaviruses in Adults: Results of 2 Double-blind, Randomized, Placebo-controlled Trials. *Clin Infect Dis*, 36:1523–1532, 2003.

[30] A. G. Porter. Picornavirus Nonstructural Proteins: Emerging Roles in Virus Replication and Inhibition of Host cell Functions. *J Virol*, 67:6917–6921, 1993.

[31] M. E. Clark, T. Hammerle, E. Wimmer, and A. Dasgupta. Poliovirus Proteinase 3C Converts an Active Form of Transcription Factor IIIC to an Inactive Form: a Mechanism for Inhibition of Host cell Polymerase III Transcription by Poliovirus. *Embo J*, 10:2941–2947, 1991.

[32] D. A. Matthews, P. A. Dragovich, S. A. Webber, S. A. Fuhrman, A. K. Patick, L. S. Zalman, T. F. Hendrickson, R. A. Love, T. J. Prins, J. T. Marakovits, R. Zhou, J. Tikhe., C. E. Ford, J. W. Meador III, R. A. Ferre, E. L. Brown, S. L. Binford, M. A. Brothers, D. M. DeLisle, and S. T. Worland. Structure-Assisted Design of Mechanism-Based Irreversible Inhibitors of Human Rhinovirus 3C Protease with Potent Antiviral Activity Against Multiple Rhinovirus Serotypes. *Proc Natl Acad Sci U S A*, 96:11000–11007, 1999.

[33] D. A. Matthews, W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, C. L. Gribskov, and Worland S. Structure of Human Rhinovirus 3C Protease Reveals a Trypsin-like Polypeptide Fold, RNA-Binding Site, and Means for Cleaving Precursor Polyprotein. *Cell*, 77:761–771, 1994.

[34] S. H. Reich, T. Johnson, M. B. Wallace, Kephart S. E., S. A. Fuhrman, S. T. Worland, D. A. Matthews, T. F. Hendrickson, F. Chan, J. W. Meador III, R. A. Ferre, E. L. Brown, D. M. DeLisle, A. K. Patick, S. L. Binford, and C. E. Ford. Substituted Benzamide Inhibitors of Human Rhinovirus 3C Protease: Structure-Based Design, Synthesis, and Biological Evaluation. *J Med Chem*, 43:1670–1683, 2000.

[35] T. O. Johnson, Y. Hua, H. T. Luu, E. L. Brown, F. Chan, S. S. Chu, P. S. Dragovich, B. W. Eastman, R. A. Ferre, S. A. Fuhrman, T. F. Hendrickson, F. C. Maldonado, D. A. Matthews, J. W. Meador III, A. K. Patick, S. H. Reich, D. J. Skalitzky, S. T. Worland, M. Yang, and L. S. Zalman. Structure-Based Design of a Parallel Synthetic Array Directed Toward the Discovery of Irreversible Inhibitors of Human Rhinovirus 3C Protease. *J Med Chem*, 45:2016–2023, 2002.

[36] M. G. Cordingley, R. B. Register, P. L. Callahan, V. M. Garskey, and R. J. Colonno. Cleavage of Small Peptides In Vitro by Human Rhinovirus 14 3C Protease Expressed in *Escherichia coli*. *J Virol*, 63:5037–5045, 1989.

[37] D. C. Orr, A. C. Long, J. Kay, B. M. Dunn, and J. M. Cameron. Hydrolysis of a Series of Synthetic Peptide Substrates by the Human Rhinovirus 14 3C Proteinase, Cloned and Expressed in *Escherichia coli*. *J Gen Virol*, 70:2931–2942, 1989.

[38] A. C. Long, D. C. Orr, J. M. Cameron, B. M. Dunn, and J. Kay. A Consensus Sequence for Substrate Hydrolysis by Rhinovirus 3C Protease. *FEBS L*, 258:75–78, 1989.

[39] F. G. Hayden, R. B. Turner, J. M. Gwaltney, K. Chi-Burris, M. Gersten, Hsyu P., A. K. Patick, G. J. Smith III, and L. S. Zalman. Phase II, randomized, double-blind, placebo-controlled studies of ruprintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers. *Antimicrob Agents and Chemo*, 47:3907–3916, 2003.

[40] M. Xian, Q. M. Wang, X. Chen, K. Wang, and P. G. Wang. S-Nitrosothiols as Novel, Reversible Inhibitors of Human Rhinovirus 3C Protease. *Bioorg and Med Chem Letters*, 10:2097–2100, 2000.

[41] S. H. Chen, J. Lamar, F. Victor, N. Snyder, R. Johnson, B. A. Heinz, M. Wakulchik, and Q. M. Wang. Synthesis and Evaluation of Tripeptidyl Alpha-Ketoamides as Human Rhinovirus 3C Protease Inhibitors. *Bioorg and Med Chem Letters*, 13:3531–3536, 2003.

[42] P. S. Dragovich, T. J. Prins, R. Zhou, E. L. Brown, F. C. Maldonado, S. A. Fuhrman, L. S. Zalman, T. Tuntland, C. A. Lee, A. K. Patick, D. A. Matthews, T. F. Hendrickson, M. B. Kosa, B. Liu, M. R. Batugo, JP. R. Gleeson, S. K. Sakata, L. Chen, M. C. Guzman, J. W. Meador III, R. A. Ferre, and S. T. Worland. Structure-Based Design, Synthesis, and Biological evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. 6. Structure-Activity Studies of Orally Bioavailable, 2-Pyridone-Containing Peptidomimetics. *J Med Chem*, 45:1607–1623, 2002.

[43] D. S. Wishart, B. D. Sykes, and F. M. Richards. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry*, 31:1647–1651, 1992.

[44] D. Sharma and K. Rajarathnam. ^{13}C NMR Chemical Shifts Can Predict Disulfide Bond Formation. *J Biomol NMR*, 18:165–171, 2000.

[45] P. Hudaky and A. Perczel. Toward Direct Determination of Conformations of Protein Building Units from Multidimensional NMR Experiments VI: Chemical Shift analysis of His to Gain 3D Structure and Protonation State Information. *J Comp Chem*, 26:1307–1317, 2005.

[46] N Tjandra and A. Bax. Solution NMR Measurement of Amide Proton Chemical Shift Anisotropy in ^{15}N-Enriched Proteins. Correlation with Hydrogen Bond Length. *J Am Chem Soc*, 119:8076–8082, 1997.

[47] L. Braunschweiler and R. R. Ernst. Corhearence Transfer by Isotropic Mixing: Application to Proton Correlation Spectroscopy. *J Magn Reson*, 53:521–528, 1983.

[48] D. R. Muhandiram and L. E. Kay. Gradient-Enhanced Triple-Resonance 3-Dimensional NMR Experiments with Improved Sensitivity. *J Magn Reson B*, 103:203–216, 1994.

[49] M. Wittekind and L. Mueller. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha and Beta-Carbon Resonances in Proteins. *J Magn Reson B*, 101:201–205, 1993.

[50] C. Griesinger, O. W. Sørensen, and R. P. Ernst. Two-dimensional Correlation of Connected NMR Transitions. *J Am Chem Soc*, 107:6394–6396, 1985.

[51] G. W. Vuister and A. Bax. Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond J(HNHα) Coupling Constants in ^{15}N-Enriched Proteins. *J Am Chem Soc*, 115:7772–7777, 1993.

[52] M. Karplus. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *J Chem Phys*, 30:11–15, 1959.

[53] D. Marion, L. E. Kay, S. W. Sparks, D. A. Torchia, and A. Bax. Three-dimensional Heteronuclear NMR of ^{15}N Labeled Proteins. *J Am Chem Soc*, 111:1515–1517, 1989.

[54] O. Zhang, L. E. Kay, J. P. Olivier, and J. D. Forman-Kay. Backbone ^{1}H and ^{15}N Resonance Assignments of the N-terminal SH3 Domain of Drk in Folded and Unfolded States Using Enhanced-Sensitivity Pulsed Field Gradient NMR Techniques. *J Biomol NMR*, 4:845–858, 1994.

[55] K. Wüthrich. *NMR of Proteins and Nucleic Acids.* Wiley Press, New York, NY, USA, 1986.

[56] B. Whitehead, C. J. Craven, and J. P. Waltho. *Protein NMR Techniques (Methods in Molecular Biology).* Humana Press, Totowa, NY, USA, 1997.

[57] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton. *Protein NMR Spectroscopy: Priciples and Practice.* Academic Press, New York, NY, USA, 1996.

[58] M. Ikura, L. E. Kay, and A. Bax. A Novel Approach for Sequential Assignment of ^{1}H, ^{13}C, and ^{15}N Spectra of Larger Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin. *Biochemistry*, 29:4659–4667, 1990.

[59] S. Grzesiek, J. Anglister, and A. Bax. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in ^{13}C/^{15}N-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. *J Magn Reson*, 101:114–119, 1993.

[60] L. E. Kay, G. Y. Xu, A. U. Singer, D.R. Muhandiram, and J. D. Forman-Kay. A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain ^{1}H and ^{13}C Correlations in H_2O Samples of Proteins. *J Magn Reson B*, 101:333–337, 1993.

[61] A. Bax, G. M. Clore, and A. M. Gronenborn. ^{1}H-^{1}HCorrelation Via Isotropic Mixing of ^{13}C Magnetization, a New 3-Dimensional Approach for Assigning ^{1}Hand ^{13}CSpectra of ^{13}C-Enriched Proteins. *J Magn Reson*, 88:425–431, 1990.

[62] K. H. Gardner, R. Konrat, M. K. Rosen, and L. E. Kay. An (H)C(CO)NH-TOCSY Pulse Scheme for Sequential Assignment of Protonated Methyl Groups in Otherwise Deuterated ^{15}N,^{13}C-labeled Proteins. *J Biomol NMR*, 8:351–356, 1997.

[63] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, Read R. J., L. M. Rice, T. Simonson, and G. L. Warren. Crystallography & NMR system: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr D Biol Crystallogr*, 54:905–921, 1998.

[64] P. Güntert, C. Mumenthaler, and K. Wüthrich. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol*, 273:283–298, 1997.

[65] A. J. Nederveen, J. F. Doreleijers, W. Vranken, Z. Miller, C. A. E. M. Spronk, S. B. Nabuurs, P. Guentert, M. Livny, J. L. Markley, M. Nilges, Ulrich E. L., R. Kaptein, and A. M. Bonvin. RECOORD: A Recalculated coordinate Database of 500+ Proteins from the PDB Using Restraints from the BioMagResBank. *Proteins*, 59:662–672, 2005.

Chapter 2

Expression, Purification and Sample Preparation of the HRV14-3C Protease

2.1 Introduction

The yield of pure protein is an important consideration in obtaining high quality NMR spectra. Likewise, as outlined in Chapter 1, the production of protein samples for NMR frequently requires the incorporation of diamagnetic isotopes (^{15}N and ^{13}C). Isotopic labelling can be costly depending on the amount of protein expressed per litre of production culture and the amount of protein lost during purification. In addition, poor sample stability necessitates the need for larger production volumes and/or increased number of production runs. This fact alone can make the study of some proteins by NMR quite prohibitive. However, sample recovery *via* refolding protocols can play a pivotal role in reducing costs and ensuring project success. In many cases, the development of such protocols may be needed if the sample precipitates out of solution after only a few days or weeks following concentration.

The complexity of a given sample preparation can significantly affect production yields. Multiple steps involving precipitation, refolding, multiple chromatography runs and dialysis complicate protocols, increase the potential for procedural errors, negatively impact experimental reproducibility and reduce the quantity of protein recovered for subsequent steps. Because no step along the production pipeline is ever 100% efficient, the availability of material for subsequent steps is dependent on the efficiency of the preceding step and the technical skills performed at the bench.

Previously published expression and purification protocols for the HRV14-3C protease describe a variety of complex multi-step procedures with a multitude of different parameters [1, 2, 3, 4]. However, due to their complexity, attempts to reproduce some of these purification protocols resulted in mixed results. Furthermore, the majority of the HRV14-3C protease studies published prior to our investigation focused on biological questions that did not require the production of large quantities of pure protein. In fact, only one paper explored the production of NMR samples [3] with reported yields of only 8.5 mg/L. Unfortunately, no additional expression or purification optimization was reported subsequent to this paper.

Therefore, it became evident that before initiating any structural work on the HRV14-3C protease, a re-examination of the expression and purification protocols for this enzyme would be required. Solution conditions that balanced the stability of the protein with sample conditions ideal for NMR experimentation would also require derivation. The significance of these concerns is highlighted by the fact that the HRV14-3C protease is a relatively large protein (182 residues) for NMR studies. The previous chapter emphasized the correlation between protein size to the quality of NMR data. This relationship played an important part in our efforts directed at optimizing the expression and purification of HRV14-3C. This chapter describes the pivotal experiments that directed the decision process in achieving the following goals:

1. Development of an efficient expression and purification protocols for the HRV14-3C protease.

2. Development of optimal sample conditions for NMR experimentation.

2.2 Materials and Methods

Sterile microbiology procedures were used throughout the expression, purification and sample preparation. Equipment and materials were either flame or autoclave-sterilized. Temperature sensitive solutions or buffers were filtered through 0.22 μm cut-off membranes. Agar plates were prepared in a laminar flow hood. All formulas and recipes used in the production and purification of HRV14-3C where adapted from *Molecular Cloning: A Laboratory Manual* [5] and are outlined in *Appendix A*.

2.2.1 SDS-PAGE Gel Analysis

Freshly prepared 12% gels (70 mm) were placed into BIORAD™ Mini-PROTEAN II cells, covered with running buffer. All trapped air bubbles subsequently removed. Samples were combined with 2x loading dye (1:1 ratio) in 1.5 ml epindorf tubes and heated above 100 °C for 5 to 10 minutes for supernatant and pelleted samples respectively. Once cool, 5 μl of each sample were pipetted into separate wells and loaded into the gel. The gel was initially run with a running voltage of 90 V for 5 minutes. Samples were separated using a voltage of 180 V for 50 to 60 minutes. Following the separation, the gels were rinsed with water and submerged in staining buffer whereby staining ensued for 1 hour on a rotating bed at 30 RPM. Following this, the staining buffer was decanted and the de-staining solution added. The destaining procedure was allowed to proceed overnight until the gel was clear and protein bands were visible. Gels were dried at 25 °C for 24 to 48 hours between layers of pre-softened cellophane stretched over a square frame after being rinsed with dd H_2O.

2.2.2 Enzymatic Activity Assay

The presence of HRV14-3C in different chromatographic fractions was confirmed with a commercially available colorimetric assay from Bachem™ (EALFQ-p-nitroaniline; product # M-2075). Reactions were conducted using:1999sing 25 μl aliquots of purified HRV14-3C protease in chromatography buffer (*Appendix A*) and 50 μl of 2.5 mM EALFQ-pNA in 750 μl of dd H_2O. This resulted in a final buffer concentration of 0.75 mM TRIS-HCl, 150 μM DTT and 62.5 μM EDTA. The reaction temperature was 25 °C. Peptide cleavage by HRV14-3C yielded free p-nitroaniline that was monitored at 405 nm [6] for 10 minutes. Positive results with this assay not only confirmed the presence of HRV14-3C but also identified which remaining fractions should be checked *via* SDS-PAGE gel analysis for HRV14-3C presence and purity.

2.2.3 Bradford Assay

The BIORAD Bradford assay [7] was used to quantify the concentrations of all HRV14-3C samples. A stock solution of BSA (1.44 mg/ml) was diluted to provide 6 standard solutions with concentrations of 0.2, 0.3, 0.4, 0.6, 0.7 and 0.9 mg/ml. Brilliant Blue G-250

dye solution was diluted 5 fold in dd H_2O and filtered through Whatman #1 filter paper. One hundred μl of each standard was added to 5 ml of dye and allowed to react for 5 minutes before measuring the absorbency at 595 nm on an Pharmacia-Biotech UltraSpec 3000 UV spectrometer. These data were fitted and the linear relationship was used to quantify the concentrations of unknown HRV14-3C samples.

2.2.4 Chromatography Resin Binding Experiments

Fifty μl aliquots of the *E. coli* lysate supernatant (described in section 2.2.6), which contained the HRV14-3C protease in lysing buffer (*Appendix A*) at pH 9.0 - 9.2, were added to various anionic and cationic Sepharose™ resins (Q, DEAE, SP and CM) pre-equilibrated in 20 mM phosphate buffer (0.5mM EDTA and 2mM DTT) with various pHs (9.3 to 4.3) at a 1:2 ratio in 1.5 ml epindorph tubes. The resulting pHs were measured and the samples were placed onto a rocker bed at room temperature for 20 minutes, which allowed the proteins to bind to the resin. Aliquots of these samples (5 μl) were then removed and run on SDS-PAGE gels to analyze protein binding.

2.2.5 HRV14-3C Protease Expression

The pET-3a plasmid containing the HRV14-3ABC gene (a gift from M.N.G. James, University of Alberta) was transfected into BL21(DE3) pLysS *E. coli* cells via electroporation (1.8 kV for 5 ms). One of the cultures from this procedure, having confirmed expression and antibiotic resistance, was isolated and used to make a 50% glycerol freezer stock (labelled OCT22-D). This stock was stored at -80 °C and used for all subsequent production of the HRV14-3C protease.

E. coli cells (OCT22-D) were streaked onto LB media plates containing ampicillin (100 μg/ml) and chloramphenicol (50 μg/ml) using a sterile platinum loop and incubated at 37 °C for 12 to 24 hours. A starter (overnight) culture was prepared by transferring a few isolated colonies from the LB plate using a flame sterilized platinum loop to 25 ml of filter sterilized TB media in a 125 ml bevel flask and capped with sterile foil. The overnight culture was agitated at 37 °C on a rotating bed at 325 RPM for 20 to 24 hours. The overnight culture was allowed to equilibrate to room temperature and used to inoculate 1 L of TB media. The 1 L production culture was inoculated with enough overnight culture to

produce an initial OD_{600} of ~0.1 (~10 ml starter culture) and separated into four 1 L bevel flasks (250 ml media per flask). These cultures were covered with sterile foil and agitated on a bench-top shaker-incubator at 325 RPM at 37 °C until the OD_{600} reached ~0.8 - 1.2. Upon reaching this optical density, production cultures were induced with 1.6 mM IPTG. The temperature was reduced to 30 °C and the cells were allowed to express the HRV14-ABC poly-protein for 12 to 14 hours. Booster doses of antibiotics were administered at 6 to 8 hour intervals throughout the culture growth and protein expression phases to maintain plasmid presence. Booster doses of chloramphenicol however, were neither required nor administered during expression as this antibiotic is required to maintain the repression of T7 RNA polymerase during production culture amplification. Protein expression protocol modifications as described by Marley, *et al.* [8] were utilized to adapt the expression protocol for [U-^{13}C/^{15}N]-HRV14-3C samples. This protocol involved initial growth of *E. coli* cells in 1 L TB media. These cells were subsequently followed by pelletint in a Sorvall RC-3 swinging bucket centrifuge at 3000 RPM (2300 x g) for 30 minutes once the cellular OD_{600} reached 0.8 to 1.0. The TB media was then decanted and the cells were rinsed with an equal amount of PBS. The PBS was removed in the same manner (*ie.* re-pelleting the cells followed by cantation of the saline). The cells were then resuspended in 250 ml of [U-^{13}C/^{15}N]-enriched minimal media[1] and returned to the rotating bed at 325 RPM and 37 °C. Following a one hour recovery phase at these conditions the cells were induced with 1.6 mM IPTG and the temperature was reduced to 25 °C. Protein production was allowed to ensue for 12 to 14 hours upon which the cells were harvested by centrifugation in a Sorvall RC-3 swinging bucket centrifuge at 3000 RPM (2300 x g) for 30 to 60 minutes. The minimal medial was decanted and the cell pellet stored at -20 °C until required for protein purification.

2.2.6 HRV14-3C Protease Purification

Following the production phase, the cells were harvested by pelleting in a Sorvall RC-3 swinging bucket centrifuge at 3000 RPM (2300 x g) for 30 to 60 minutes. The supernatant was decanted off and the pelleted cells were resuspended in 40 ml lysing buffer. Cell lysis

[1] Isotopic labelling was achieved by replacing NH_4Cl and glucose listed in the MM recipe (*Appendix A*) with ^{15}N-enriched NH_4Cl and ^{13}C-enriched glucose variants.

was performed with three freeze-thaw cycles (-70 to 4 °C). Thirty to forty drops of 10% PEE was added and gently mixed into the lysed cells to precipitate DNA. The lysed cells were then centrifuged at 15,000 RPM (26,000 x g) for 30 to 40 minutes to pellet the cellular and precipitated matter. The resulting supernatant was collected *via* pipette and layered onto a chromatography column filled with either Q-Sepharose™ or DEAE-Sepharose™ anionic exchange resin pre-equilibrated at 4 °C with chromatography buffer (*Appendix A*). Eight to nine millilitre fractions were collected at a rate of 3 to 5 ml/min. The HRV14-3C protein eluted shortly after the initial flow-through with greater than 90% purity (Figure 2.1). Detection of the HRV14-3C protease was achieved with the enzymatic assay (section 2.2.2) and its purity was confirmed with SDS-PAGE gel analysis (section 2.2.1). Fractions containing the HRV14-3C protease were loaded to a 50 ml BioGel-HA® hydroxyapatite column[2] and eluted with a 40 - 175 mM gradient of K_2HPO_4. Pure HRV14-3C fractions were pooled and dialyzed in phosphate buffer (*Appendix A*) using 5000 MW cut-off dialysis membranes. Quantification of HRV14-3C was conducted using the Bradford assay [7]. Samples obtained with the methods outlined here provided yields that exceeded 60 mg/L.

2.2.7 Button Test

Hampton Research microdialysis cells (10 μl) were used for screening NMR sample conditions. A purified sample of HRV14-3C (~0.12 mM) was first dialyzed into dd H_2O at pH 7.0 and allowed to equilibrate to 25 °C. A 12 μl sample was placed into each microdialysis button and air bubbles removed. The microdialysis buttons containing the sample were covered with 5000 MW cut-off dialysis membrane, which was secured with an O-ring, and placed in 10 ml of 20 mM potassium phosphate buffer with various pHs (4.7 - 7.5) and NaCl concentrations (0 - 150 mM) and allowed to equilibrate and exchange buffer at 25 °C. Sterile procedures were employed throughout the experiment to prevent sample contamination. The conditioning experiment was allowed to proceed for 3 weeks. Samples were checked for precipitation using a WILD M20 microscope at 300x magnification.

[2]This step was only employed for the production of the [U-$^{13}C/^{15}N$]-apo HRV14-3C NMR sample used to collect data subsequent to the backbone assignment experiments.

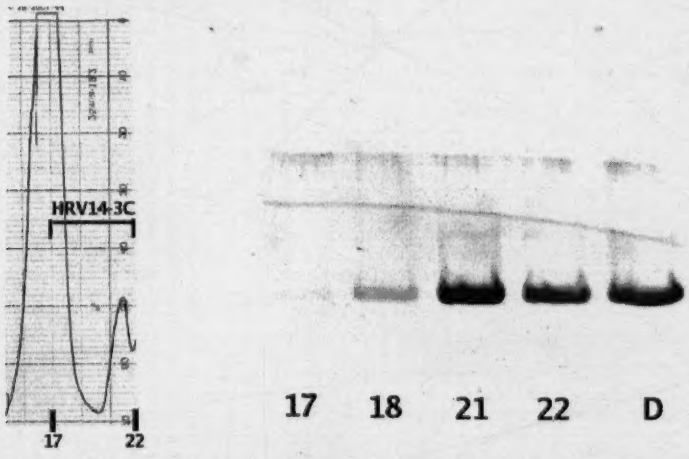

Figure 2.1: **DEAE-Sepharose™ Chromatography Results.** The chromatography UV absorbency trace measured at 280 nm and the corresponding SDS-PAGE gel fractions are shown. HRV14-3C eluted in fractions 18 to 22, which are marked on the chromatography trace. The lack of UV absorption is due to the lack of tryptophan residues in HRV14-3C's primary sequence. The lane marked D contains the pooled fractions containing the pure HRV14-3C protease after dialysis.

2.2.8 Refolding Procedure

A convenient method to recover the precipitated protease was derived from a variety of reported purification schemes that employed protein refolding subsequent to denaturing with 7 M urea [3]. Adaptations to these protocols provided a method that recovered in excess of 80% active protease (confirmed with the colorimetric EALFQ-pNA assay [6]). The devised protocol employed the use of 20 mM cysteine along with DTT as reducing agents. The DTT concentration was in excess of 10-fold that of the protease. Precipitated HRV14-3C samples were diluted to a concentration below 1 mg/ml in denaturing buffer. Solubilized protein samples were transferred to 5000 MW cut-off dialysis membrane and dialyzed against 20 mM phosphate buffer (pH 6.5, 0.5 mM EDTA and 15 mM DTT). Dialysis was carried out at 4 °C in 1 L graduated cylinders with magnetic stir bars. The buffer was exchanged every 4 to 12 hours in which the buffer was changed every 4 to 6 hours during the initial 3 exchanges to reduce the concentration of urea as quickly as possible. A minimum of 6 buffer changes were performed to ensure urea concentrations decreased to sub nM levels.

2.2.9 NMR Sample Preparation

HRV14-3C protease samples were concentrated to 0.5 - 1.0 mM *via* ultrafiltration. D_2O (10%) was added to maintain the spectrum lock and 0.1 mM DSS was added for chemical shift referencing[3]. All samples were filtered through Millipore Ultrafree-MC® 0.45 μm cut-off filters. Protein samples were transferred to either Wilmad® 5 mm thin-walled NMR tubes and sealed under argon or to SHIGEMI® microcell NMR tubes.

2.3 Results and Discussion

2.3.1 HRV14-3C Protease Expression

Upon beginning the HRV14-3C project, some initial milestones had been met by other researchers. The HRV14-3ABC gene construct had been inserted into the pET-3a plasmid between the BamHI restriction endonuclease sites and this vector had been transformed into *E. coli* competent cells (BL21(DE3) pLysS) *via* electroporation (1.8 kV for 5 ms). Upon expression, the 3C protease self-cleaves from the 3ABC gene product. Confirmation of this was reported by Birch and coworkers [3] and confirmed by SDS-PAGE gel analysis. Initial attempts at purifying the protease provided sufficient yields (~10 mg/L), to confirm the presence of active protease using the EALFQ-p-nitro-aniline colorimetric assay [6] available from Bachem™. However, the amount of soluble protein in the supernatant 6 hours after induction was relatively low in comparison to the whole cell extract and the purity was far below acceptable levels for initiating NMR experimentation. Attempts to retrieve soluble protein from the precipitated pelleted involved a variety of techniques that included dissolving *via* sonication, denaturation with urea and subsequent refolding *via* dialysis [9], selective precipitation with $(NH_4)_2SO_4$ and ion exchange chromatography. Many of these protocols were explored previously by other groups [1, 2, 3]. Unfortunately, attempts to replicate many of these protocols resulted in failure. Additionally, work conducted by Birch and coworkers [3] identified the majority of this insoluble pelleted material to be comprised of truncated HRV14-3C protease.

[3] 0.1 mM NaN_3 was added to the [U-^{13}C/^{15}N]-apo HRV14-3C NMR sample used to collect data subsequent to the backbone assignment experiments.

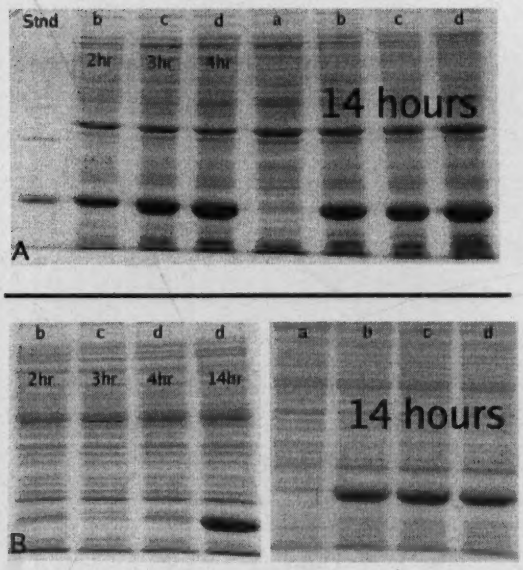

Figure 2.2: **SDS-PAGE Gel: Induction Experiments.** Gel **A** shows pelleted fractions from samples taken at various times following induction. Gel **B** shows the corresponding supernatant fractions. Lower case letters denote the various samples. Lane **a** is the uninduced control sample. The standard lane (**Stnd**) contains relatively pure HRV14-3C protease obtained from a previous expression run.

Given this insight, our attention was redirected at maximizing the amount of soluble expressed protein, rather than recovering possible insoluble, truncated material. The yield of HRV14-3C protease was successfully maximized by lowering the expression temperature to room temperature (~25 °C) from the original expression temperatures of 30 °C [10] following the initial cell growth at 37 °C. Yields were also improved by increasing the expression time from 6 to 14 hours. These results are shown in Figure 2.2. As seen from this figure, the amount of material in the pellet increases with the length of time post induction, while the amount of soluble material remains non-existent for up to 4 hours following induction. After this time however, the relative amount of precipitated material does not seem to increase while the amount of soluble HRV14-3C does. Fourteen hours after induction there is a 60:40 split between precipitated protease in the pellet and soluble protease in the supernatant. The amount of protein in the soluble fraction is consistent with reports by Knott and coworkers [1]. These yields were further increased with administration of ampicillin (100 μg/ml) after 6 hours of induction to maintain the presence of the pET-3a HRV14-3ABC plasmid.

N164D deamidation products have been previously reported with the HRV14-3C protease [11]. The production of these products could result from host cell or auto-catalytic events because the asparagine residue is followed by a glycine and because the cell lysis was conducted at alkaline pH. To limit potential enzyme mediated deamidation, all the protein purification steps were conducted at 4 °C and thermal cycling was carried out from -70 to

4 °C. These temperatures prevented the effective use of DNAse and contributed substantially to the purification time. However, this approach significantly added to the fraction of soluble protein obtained.

The production of protein NMR samples generally requires ^{13}C and ^{15}N isotopic labelling. Although labelled yeast and algae cell extracts have been employed with success [12], the expression in minimal media (MM) with single sources of these isotopes (^{13}C-glucose and ^{15}N-NH_4Cl) is a cost-effective alternative. Initially, to incorporate these isotopes in minimal media, the *E. coli* BL21(DE3) pLysS cells were conditioned by growing them on MM plates. However, it was later observed that the Marley protocol [8] eliminated the need for MM prior to expression and provided good yields of pure protease (~60mg/L). This protocol was used later in the production of all [U-^{13}C/^{15}N]-HRV14-3C protease samples subsequent to the preparation of the first double-labelled sample.

2.3.2 HRV14-3C Protease Purification

Figure 2.3 shows the resulting cell extract supernatant when applied to DEAE-Sepharose™ beads. The systematic reduction of the pH produced an increasingly cationic state of the HRV14-3C protease (theoretical pI 8.45[4] calculated with ProtParam [13]), which resulted in reduced binding to the anion exchange resin. As the pH becomes more alkaline and the protein becomes more anionic there is binding of the HRV14-3C protease along with the majority of *E. coli* proteins (above pH 7.25). Isoelectric focussing work was never conducted, however, it was hypothesized that the apparent binding of HRV14-3C above pH 7.25 was possibly a N164D deamidated form. Not readily apparent from this SDS-Page gel is

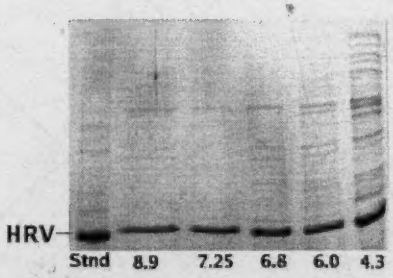

Figure 2.3: **Sepharose™ Chromatography Resin Binding Experiment.** Various supernatant samples following application to DEAE-Sepharose™ beads at different pHs, which are displayed at the bottom of each lane.

[4]Isoelectric focusing work done by Cox *et al.* [11] confirmed the isoelectric point (pI) of the native HRV14-3C to actually be 8.9 and the pI of the N164D deamidation product to be 8.3.

that the binding of HRV14-3C was not proportional to the contaminant *E. coli* proteins. In particular, the contaminant proteins bound with greater affinity due to their lower pIs compared with the HRV14-3C protease. These results were used by Birch and co-workers [3] to extract HRV14-3C granules at pH 7.5, which then required further ion exchange chromatography and refolding steps. This strategy was also utilized in the protocol described here (*Appendix B*) whereby an anionic exchange resin was used as a filtration step. However, the supernatant containing the native HRV14-3C protease was applied to the resin near HRV14-3C's isoelectric point. These results are shown in Figure 2.1 where the supernatant and the DEAE-Sepharose™ anion exchange column had their pHs adjusted to the protein's pI (~8.8 - 9.2). The HRV14-3C protease eluted just after the initial flow-through with greater than 90% purity. These results were repeated when the supernatant was subjected to a stronger anionic binding resin, Q-Sepharose™ (Figure 2.4).

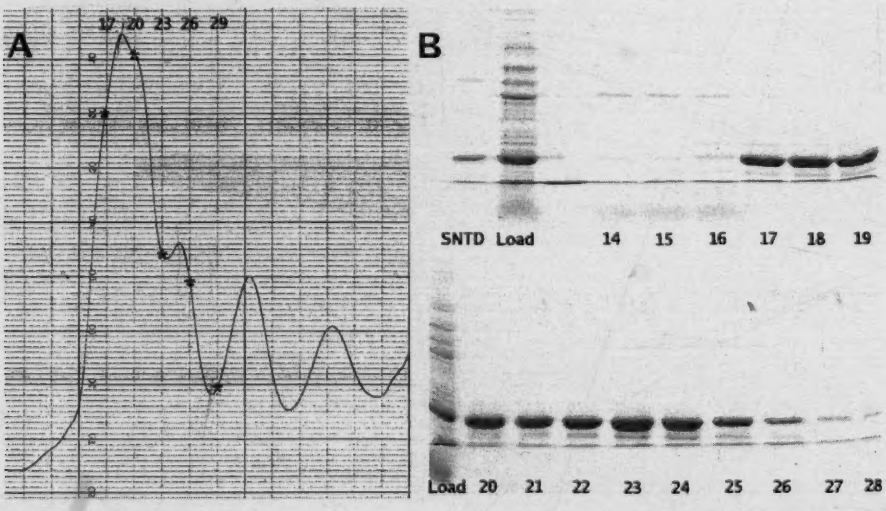

Figure 2.4: **Q-Sepharose™ Purification of the HRV14-3C Protease.** Trace A shows the UV absorbency at 280 nm for HRV14-3C purification fractions run through a 120 ml Q-Sepharose™ chromatography column at 4 °C (pH 9.0; flow rate 2.5 ml/min). Fraction numbers are give at the top of the trace with the corresponding positions marked on the trace with an asterisk. The SDS-PAGE Gels show the corresponding purification results for fractions collected in Trace A. The weak binding of the HRV14-3C protease around pH 9.0 to the anionic exchange column affords a slightly longer retention time compared with contaminant *E. coli* proteins and provides the elution of pure protease.

These data show the HRV14-3C protease eluting through the column just after the con-

taminant *E. coli* proteins with higher pIs. The weak binding of the HRV14-3C protease with the anionic exchange column led to a slightly longer retention time compared with the contaminant proteins and yielded very pure protease.

An additional step following the Q-Sepharose™ column was employed for the purification of a [U-^{13}C/^{15}N]-apo HRV14-3C protease sample used for collecting NMR spectra subsequent to the backbone assignments [14]. This step involved binding the protease to a 50 ml BioGel-HA® hydroxyapatite column. The protease was eluted with a 40 to 175 mM potassium phosphate gradient wash using a flow rate of 2 to 3 ml/min and concentrated to 0.99 mM in the NMR sample buffer with the addition of 0.1 mM NaN_3. These changes afforded a dramatic increase in sample stability from ~3 days to ~6 months. The exact reason for this enhanced stability remains unclear and was not investigated. Although higher sample purity was achieved with the use of an additional chromatography column, the inclusion of NaN_3 may have provided non-specific inhibition of the enzyme as azides have been used as moieties for drug development against the 3C proteases [15, 16, 17].

2.3.3 HRV14-3C Chromatography Fraction Detection

The HRV14-3C protease lacks a UV absorption band at 280 nm (UV_{280nm}) due to the absence of tryptophan residues in its amino acid sequence.

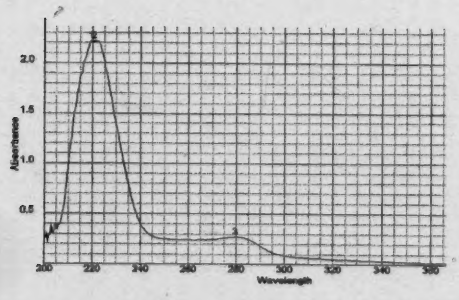

Figure 2.5: **$UV_{200-360nm}$ Scan of the HRV14-3C Protease.** The absorption UV trace from 200 to 360 nm for the pooled dialysis fraction shown in Figure 2.1.

A pure sample of the HRV14-3C protease (the dialyzed fraction shown in Figure 2.1, lane D) was subjected to a UV absorption scan (Figure 2.5), which confirmed the poor absorption at 280 nm. Consequently, detecting its elution from the chromatography column was hindered. These results are illustrated in both Figures 2.1 and 2.4, where the elution trace does not correlate with the corresponding fractions containing HRV14-3C as confirmed with SDS-PAGE gel analysis. Therefore, the p-nitroaniline assay (section 2.2.2) proved important for identifying which

chromatography fractions might contain the HRV14-3C protease. The time of HRV14-3C elution was estimated from column volume and flow rate measurements and used to test the suspected fractions for active HRV14-3C protease with the pNA activity assay. Subsequent to identifying the first fraction containing the HRV14-3C protease, consecutive fractions were subjected to SDS-PAGE gel analysis to confirm both the presence and purity of the HRV14-3C protease.

2.3.4 Protease Quantification

The low UV_{280} absorption also led to errors when quantifying the HRV14-3C protease when using methods relying on predicted extinction coefficients alone. Protein concentration can normally be estimated using the Beer-Lambert equation:

$$C * \varepsilon * l = ABS_{280nm} \tag{2.1}$$

Where C is the protein concentration given in mg/ml, ABS_{280nm} is the UV absorption at 280 nm in units ml/mM*cm , l is the length of the light path through the sample in cm and ε is the wavelength-dependent molar absorptivity (extinction co-efficient) with units of $cm^{-1}M^{-1}$. ε, however, is calculated from the number of tryptophan and tyrosine residues (assuming the measured wavelength is 280 nm) using the equation:

$$\varepsilon = \frac{(5700 * \#W + 1300 * \#Y)}{Protein\ Mwt\ (Da)} \tag{2.2}$$

The lack of tryptophan residues in the HRV14-3C protease's primary amino acid sequence estimated it's ε around 0.26 $cm^{-1}M^{-1}$ and resulted in overinflated quantifications of protein concentration, presumably from an ε below 1 $cm^{-1}M^{-1}$ and possible UV_{280nm} absorption from contaminant sources.

Protein quantification thus relied on the use of the Bradford Assay [7], which quantifies protein samples from 5 to 100 μg by comparing the unknown samples absorption at 595 nm (ABS_{595nm}) to a standard curve created from samples with known protein concentrations. After protein samples are incubated for 5 minutes in coomassie Brilliant Blue G-250 dye, the absorbence maximum shifts from 465 nm to 595 nm when binding to arginine and aromatic residues occurs, which then stabilizes the single proton anionic form of the dye [18].

2.3.5 NMR Sample Stability

Preliminary NMR samples were prepared in slightly alkaline conditions (pH 7.5) with 100 mM NaCl and only 0.2 mM of protein. These samples exhibited poor stability and precipitated readily following concentration. The button test (Table 2.1) described by Bagby *et al.* [19] was performed to examine the stability of the HRV14-3C protease at room temperature under various pH and salt conditions.

Table 2.1: **Button Test Results for the HRV14-3C Protease after 3 weeks**

pH / [NaCl]	0 mM	50 mM	100 mM	150 mM
4.7	**	**	**	**
5.5	**	**	**	**
6.5	v	***	****	nd
7.0	*	*	*	nd
7.5	*	**	**	nd

* Indicates the relative amount of observable precipitation
v Conditions void of visible precipitation
nd Not determined due to experimental flaw

These results demonstrate that slightly acidic conditions (pH 6.5) and no salt would maintain soluble HRV14-3C protease for long periods of time (>3 weeks). These conditions closely resemble those reported by Birch *et al.* [3] who also reported the preparation of an NMR sample with the following conditions: 20 mM MES, 1 mM EDTA, 1 mM DTT, pH 6.5. Conveniently, our derived conditions (0 mM NaCl, 0.5 mM EDTA, 15 mM DTT, 20 mM KH_2PO_4 at pH 6.5) were amenable to collecting NMR experiments and were subsequently used for all HRV14-3C protein NMR sample preparations. These conditions allowed NMR samples to remain in solution at concentrations above 0.5 mM for a few hours before visible precipitation.

Both oxidized and reduced states of the protease were studied for their stability. The non-reactive, oxidized state of the protease precipitated upon oxidation *via* pumping filtered air into a dialysis vessel containing dilute HRV14-3C protease at the optimized sample conditions. In contrast, maintaining a reduced state of HRV14-3C with a buffer containing

15 mM DTT afforded soluble protease at NMR concentrations above 0.5 mM for a few days before precipitation was visible. These experiments resulted in the NMR sample conditions listed above and afforded protein stability long enough to collect a sufficient number of three dimensional heteronuclear NMR experiments. These experiments ultimately provided the backbone assignments of the apo HRV14-3C protease [14]. An initial 1D ^{1}H NMR spectrum of the apo HRV14-3C protease prepared using the outlined conditions is shown in Figure 2.6.

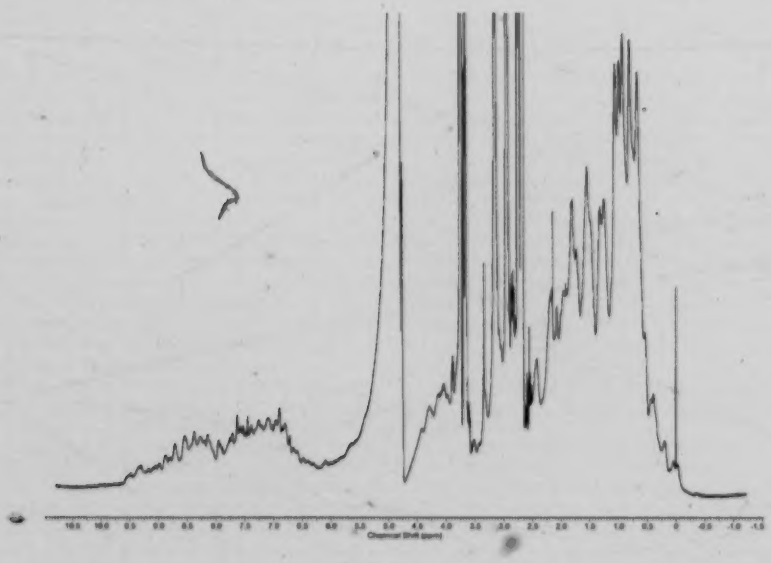

Figure 2.6: **Preliminary 1D ^{1}H NMR Spectrum of the HRV14-3C Protease.** This complex spectrum results from the Fourier transformation of the FID shown in Figure 1.5. The spectrum is referenced to DSS at 0.00 ppm.

2.4 Conclusion

The objectives for this phase of the HRV14-3C project were to obtain pure, concentrated protease with conditions that afforded protein stability long enough to collect an average three dimensional NMR experiment. In completing the experiments outlined in this chapter, the expression was optimized by minimizing the extent of inclusion body formation by lowering the expression temperature and increasing the initial time for expression time

2-fold. The purification of the HRV14-3C protease was optimized by applying the cell lysate to a single anionic exchange column. Sample stability of the apo HRV14-3C protease was further improved with the use of a second binding and elution step through a hypoxyapatite column. These protocols employ a simplified strategy and produce higher yields than any other protocol published to date. A quantification strategy that relied solely on the Bradford assay was used to compensate for the fact that the protein lacked any tryptophan residues. The initial one dimensional NMR spectra for the HRV14-3C protease (Figure 2.6) provided sufficient results to initiate sample conditioning experiments (Table 2.1). These experiments provided improved sample stability and allowed the collection of 3D heteronuclear NMR experiments. Futhermore, a refolding protocol was derived for recovering denatured protease. The exact protocols developed during this work are outlined in *Appendix B*.

Bibliography

[1] J. A. Knott, D. C. Orr, D. S. Montgomery, C. A. Sullivan, and A. Weston. The Expression and Purification of Human Rhinovirus Protease 3C. *Euro J Biochem*, 182:547–55, 1989.

[2] M. G. Cordingley, R. B. Register, P. L. Callahan, V. M. Garskey, and R. J. Colonno. Cleavage of Small Peptides In Vitro by Human Rhinovirus 14 3C Protease Expressed in *Escherichia coli*. *J Virol*, 63:5037–5045, 1989.

[3] G. M. Birch, T. Black, S. K. Malcolm, M. T. Lai, R. E. Zimmerman, and S. R. Jaskunas. Purification of Recombinant Human Rhinovirus-14 3C Protease Expressed in *Escherichia coli*. *Protein Expres Purif*, 6:609–618, 1995.

[4] G. J. Davis, Q. M. Wang, G. A. Cox, R. B. Johnson, M. Wakulchik, C. A. Dotson, and E.C. Villarreal. Expression and Purification of Recombinant Rhinovirus 14 3CD Proteinase and Its Comparison to the 3C Proteinase. *Archives Biochem and Biophys*, 346:125–130, 1997.

[5] S. Sambrook, E. F. Fritch, and T. Maniatis. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York, NY, 1989.

[6] Q. M. Wang, R. B. Johnson, G. A. Cox, E. C. Villarreal, and R. J. Loncharich. A Continuous Colorimetric Assay for Rhinovirus-14 3C Protease Using Peptide p-Nitroanilides as Substrates. *Anal Biochem*, 252:238–245, 1997.

[7] M. M. Bradford. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal Biochem*, 72:248–254, 1976.

[8] J. Marley, M. Lu, and C. Bracken. A Method for Efficient Isotopic Labeling of Recombinant Proteins. *J Biomol NMR*, 20:71–75, 2001.

[9] H. Lilie, E. Schwarz, and R. Rudolph. Advances in Refolding of Proteins Produced in *E. coli*. *Curr Opin Biotech*, 9:497–501, 1998.

[10] M. Piatak, J. A. Lane, W. Laird, M. J. Bjorn, A. Wang, and M. Williams. Expression of Soluble and Fully Functional Ricin A Chain in *Escherichia coli* is Temperature-Sensitive. *J Biol Chem*, 263:4837–4843, 1988.

[11] G. A. Cox, R. B. Johnson, J. A. Cook, M. Wakulchik, M. G. Johnson, E. C. Villarreal, and Q. M. Wang. Identification and Characterization of Human Rhinovirus-14 3C Protease Deamidation Isoform. *J Biol Chem*, 274:13211–13216, 1999.

[12] E. R. P. Zuiderweg and S. W. Fesik. Heteronuclear Three-Dimensional NMR Spectroscopy of the Inflammatory Protein C5a. *Biochemistry*, 28:2387–2391, 1989.

[13] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvarud, M. R. Wilkins, R. D. Appel, and Bairock A. *Protein Identification and Analysis Tools on the ExPASy Server: The Protomics Protocols Handbook.* Humana Press, Totowa, NJ, USA, 2005.

[14] T. C. Bjorndahl, H. Monzavi, and D. S. Wishart. Letter to the Editor: Backbone ^{1}H, ^{15}N and ^{13}C Assignments for the Human Rhinovirus 3C Protease (serotype 14). *J Biomol NMR*, 26:85–86, 2003.

[15] S. Venkatraman, J. Kong, S. Nimkar, and Q. M. Wang. Design, Synthesis and Evaluation of Azapeptides as Substrates and Inhibitors for Human Rhinovirus 3C Protease. *Bioorg and Med Chem Letters*, 9:577–480, 1999.

[16] R. D. Hill and J. C. Vederas. Azodicarboxamides: A New Class of Cysteine Proteinase Inhibitor for Hepatitis A Virus and Human Rhinovirus 3C Enzymes. 64:9538–9546, 1999.

[17] Y. Huang, B. Malcolm, and J. C. Vederas. Synthesis and Testing of Azaglutamine Derivatives as Inhibitors of Hepatitis A Virus (HAV) 3C Proteinase. *Bioorg and Med Chem*, 7:607–619, 1999.

[18] S. J. Compton and C. G. Jones. Mechanism of Dye Response and Interference in the Bradford Protein Assay. *Anal Biochem*, 151:369–374, 1985.

[19] S. Bagby, K. I. Tong, D. Kiu, J. R. Alattia, and M. Ikura. The Button Test: A Small Scale Method Using Microdialysis Cells for Assessing Protein Solubility at Concentrations Suitable for NMR. *J Biomol NMR*, 10:279–282, 1997.

Chapter 3

Synthesis and Evaluation of the HRV14-3C Protease Inhibitor: Acetyl-LEALFQ-Ethyl Propenoate

3.1 Introduction

The intent of this study was to complete the apo structure of the HRV14-3C protease. However, the initial structural work conducted on the uninhibited protease using standard NMR methods presented some difficulties that resulted from sample precipitation, possible degradation signals and a number of missing amide signals in the ^{15}N-HSQC spectrum. At this time, some interesting correlations were also made from homologous 3C proteases. Specifically, the HRV2-3C enzyme (51% sequence identity with HRV14-3C) was solved with a peptide-mimic inhibitor (AG7088) and presented uniformly low B-factors throughout protease, which included the substrate binding site [1]. This was in contrast to the large B-factors reported for the apo Polio-3C protease [2], which shares 49% sequence identity with the HRV14-3C protease. These large B-factors were seen in both the substrate binding site and in areas involved with interdomain contact (Figure 3.1, Figure 3.2). Therefore, it was suspected that the lack of amide signals for the HRV14-3C enzyme belonged to residues undergoing intermediate chemical exchange and these residues were localized to the proteolytic substrate binding region. Furthermore, it was hypothesized that by inactivating the enzyme, stabilization of the HRV14-3C protease might follow and these NMR signals might become visible. Indeed, if these missing signals were localized to the proteolytic site, then this re-direction would prove to be critical.

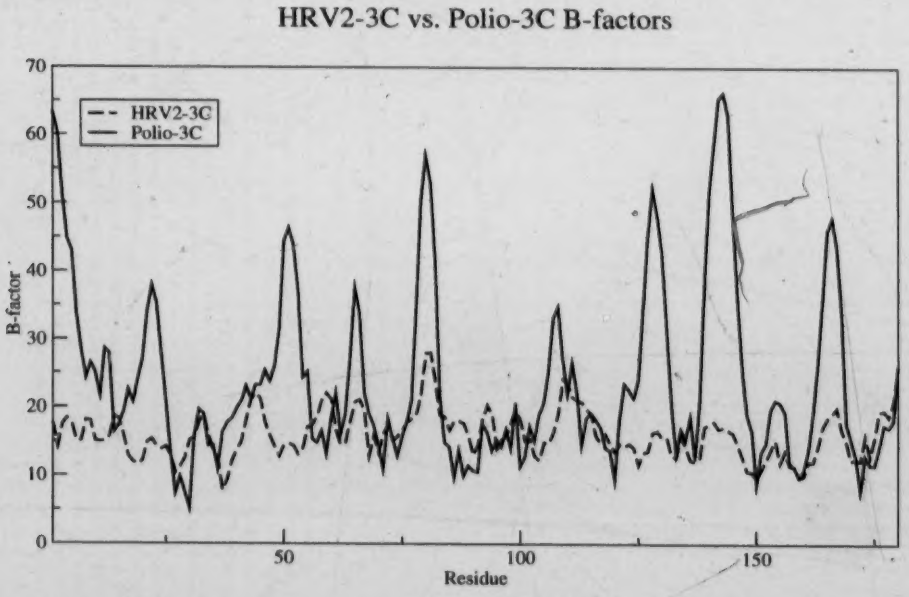

Figure 3.1: **HRV2-3C vs. Polio-3C B-factors**. Regions with large B-factors for the apo Polio-3C protease (1L1N.pdb: structure A) are localized around the proteolytic binding site and interdomain contact areas. These regions are comparably immobile in the inhibitor bound HRV2-3C protease (1QCC.pdb), which presents relatively constant B-factors throughout the structure. The B-factors reported were obtained from the backbone $^{13}C_{\alpha}$ atoms.

Without NMR data from the proteolytic residues, an assessment of the active site pharmacophore did not seem possible. It was also suspected that stability issues might also be resolved as various proteolytic cocktails are available for reducing host cell degradation of expressed proteins and improving the stability of proteolytic enzymes. Furthermore, because a number of picornaviral 3C proteases had already been studied successfully with bound inhibitors [3, 4], it was suspected that inhibition of the HRV14-3C protease with a substrate might also prove beneficial. This rationale led to the synthesis of the acetyl-LEALFQ-ethyl propenoate inhibitor that is discussed in this chapter.

Specifically, to investigate the inhibitor-bound structure of the HRV14-3C protease, an irreversible peptidyl inhibitor was designed on the basis of the modified 2C/3A cleavage sequence described by Wang *et al.* [5] and the Michael Acceptor ethyl propenoate inhibitor described by Dragovich *et al.* [6] and Matthews *et al.* [1]. The strategy reported here

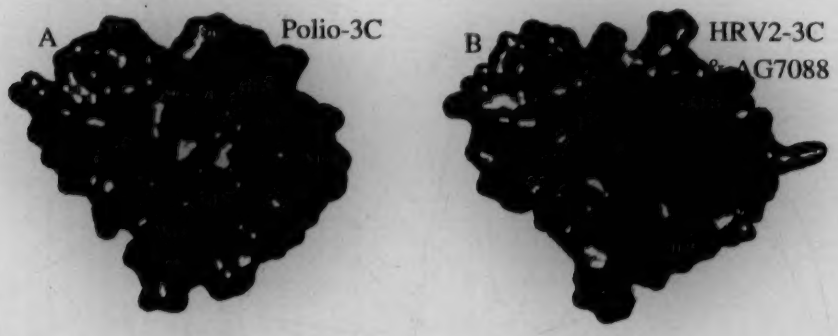

Figure 3.2: **Structural Relationship of Flexible Residues in Picornaviridae 3C Proteases**. Molecular surface renderings for the apo Polio-3C and inhibited HRV2-3C proteases. **A**: Residues with high B-factors for the apo Polio-3C protease are shown. Residues with B-factors greater than 30 are colored light blue. Residues with B-factors greater than 40 are colored dark blue. **B**: The surface rendering of the AG7088 inhibited HRV2-3C protease is shown with covalent inhibitor, AG7088 (colored red). Residues with only backbone contacts with AG7088 are colored magenta. Residues with side-chain van der Waals contacts to AG7088 are shown in blue.

builds upon previously published work that developed peptide-mimic covalent modifiers for the enzyme. However, this effort also focused on the incorporation of a relatively large natural peptide fragment of sufficient size that would allow investigation of downstream substrate interactions that have not yet been characterized. This chapter describes the synthesis of the peptidyl inhibitor used to inactivate the HRV14-3C protease. Furthermore, the characterization of the inhibitor pre-incubation with HRV14-3C is described along with the experiments conducted to confirm covalent attachment.

3.2 Materials and Methods

3.2.1 Inhibitor Synthesis

The six residue peptidyl-ethyl propenoate inhibitor was synthesized in two steps. Briefly, an acetylated peptide (Acetyl-L-Leu-L-(OtBu-Glu)-L-Ala-L-Leu-L-Phe-COO^-) was prepared via Fmoc solid phase peptide synthesis and chemically coupled in solution to a L-[(Trt)-Gln]ethyl propenoate moiety that was synthesized based on methods described by Dragovich [6]. The synthesis protocol is outlined in Figure 3.3 and the specific details follow:

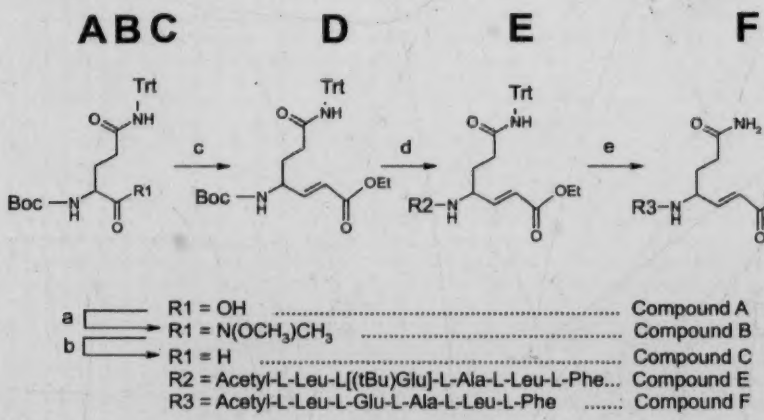

Figure 3.3: **Synthesis of the Acetyl-LEALFQ-Ethylpropenoate Inhibitor**: Reagents and conditions (Trt = CPh_3). (a) 1 mEq Isobutyl Chloroformate, 2 eq NMM, 20 minutes, 0 °C $\longrightarrow$ 1 mEq HCl $HN(OCH_3)CH_3$, 0 °C, 20 minutes $\longrightarrow$ 2 hours, 25 °C, 70%. (b) THF, 1.75 mEq DIBAL, -78 °C, 4 hours, crude. (c) THF, 1 mEq $(EtO)_2POCH_2CO_2Et$, 1 mEq $NaN(TMS)_2$, -78 °C then 20 minutes 0 °C. Crude **C** in THF, 2 hours, -78 °C $\longrightarrow$ 10 minutes, 0 °C, 45%. (d) **D** in solution with 4 M HCl in 1,4-dioxane, 3 hours, 25 °C. Then 0.21 mmol peptide, 0.27 mmol HOBt, 90 μl NMM, 0.28 mmol EDC, 24 hours, 25 °C, 50%. (e) 0.278 mmol E, 140 μl TIS, TFA, 30 minutes, 25 °C, 99%.

Acetyl-L-Leu-L-(OtBu-Glu)-L-Ala-L-Leu-L-Phe-COO^- (R2). All solid phase peptide synthesis reactions were done under normal atmospheric pressure at 25 °C in a frit plugged polyethylene column reaction vessel. The resin was washed between reactions with successive rinses and filtering of DMF (3 x 2 minutes), isopropyl alcohol (3 x 2 minutes) and DMF (3 x 2 minutes). 2-Chlorotrityl chloride resin (0.175 g, 0.1 mmol) was swollen by shaking in 2 ml of DMF for 5 minutes and filtered. Fmoc de-protection of amino acids (0.2 mmol for phenylalanine and 0.8 mmol for successive amino acids leucine, alanine, O-tButyl-glutamic acid and leucine) was accomplished by shaking the resin-bound peptide with freshly prepared 25% piperidine in DMF (5 ml). Activation of successive amino acids was accomplished by addition of N-methylmorphiline (200 μl NEET) in 2 ml DMF. The coupling agent HBTU (1.6 ml of 0.45 M solution in DMF) was used for all reactions following the coupling of phenylalanine to the 2-Chlortrityl resin. The coupling reaction mixture was agitated and allowed to proceed for 45 minutes upon which the resin was washed and small

samples were removed for the Kaiser ninhydrin test to confirm reaction completion prior to addition of subsequent amino acids. The peptide was capped with activated acetic acid (45.6 μl, 0.8 mmol) for the final coupling reaction. Peptide cleavage from the 2-Chlortrityl resin was achieved by agitation in 10 ml of acetic acid - TFA - DCM (1:1:3) for 2 hours. The peptide was rinsed with acetic acid - DCM (1:4; 2 x 10 ml) and concentrated to a white foam (130 mg, 0.189 mmol, 94%).

[Boc-L-(Trt-Gln)]-N(OMe)Me (Compound B). N-α-Boc-γ-trityl-L-glutamine, **Compound A** (1 mEq, 1.87 g, 3.67 mmol) and NMM (2 mEq, 810 μl, 7.35 mmol) were added to 25 ml DCM at 0 °C. Isobutyl chloroformate (1 mEq, 477 μl, 3.68 mmol) was added to the solution and stirred for 20 minutes. N,O-dimethylhydroxylamine hydrochloride (1 mEq, 3.67 mmol, 360 mg) was added and the solution was stirred for a further 20 minutes at 0 °C then for 2 hours at 25 °C. The reaction mixture was partitioned between H_2O (15 ml) and DCM (2x15 ml). The organic layers were combined and dried over Na_2SO_4, concentrated to 20 ml and purified over a 200 ml flash chromatography column (50-35% hexanes in EtOAc), which provided **Compound B** (1.36 g, 2.56 mmol, 70%) as a white foam.

[Boc-L-(Trt-Gln)]-H (Compound C). **Compound B** (1.36 g, 2.56 mmol, 0.70 mEq) was solubilized in 8 ml THF at -78 °C and stirred for 20 minutes. DIBAL (6.4 ml of 1 M solution in toluene, 1.75 mEq) was added and stirred at -78 °C for 4 hours. The reaction was quenched with successive additions of methanol (338.4 μl) and 1 M HCl (846 μl) and warmed to 25 °C. The suspension was diluted with ethanol (12.69 ml), washed with 1 M HCl (3 x 10 ml), 50% saturated aqueous $NaHCO_3$ (10 ml) and H_2O (10 ml). The organic layer was dried over $MgSO_4$, filtered and concentrated to give **Compound C** as a white solid.

[Ethyl 3-Boc-L-(Trt-Gln)]-(E/Z)-propenoate (Compound D). A solution of triethylphosphonoacetate (557.6 mg, 2.48 mmol, 1 mEq) in 200 μl THF was prepared at -78 °C. Sodium bis(trimethylsilyl)amide (1 mEq, 0.023 mmol, 23 μl of 1 M solution in THF) was added and stirred for 20 minutes. **Compound C** was dissolved in 10 ml THF at -78 °C and added to the reaction mixture by cannula. The reaction was stirred at -78 °C for 2 hours and warmed to 0 °C for 10 minutes. The reaction was partitioned between 0.5M HCl (7.5 ml) and 50% EtOAc in hexanes (2 x 7.5 ml). The organic layers were combined, dried over Na_2SO_4 and concentrated. The compound was purified by flash column chromatogra-

phy (50% ethyl acetate in hexanes) and concentrated to a white foam (600 mg, 1.11 mmol, 45% total yield: 60% trans , J = 17.64 Hz and 40% cis, J = 11.57 Hz).

Ethyl 3-[Acetyl-L-Leu-L-(OtBu-Glu)-L-Ala-L-Leu-L-Phe-L-(Trt-Gln)]-(E/Z)-propenoate (Compound E). **Compound D** (100 mg, 0.19 mmol) was added to a solution of 4 M HCl in 1,4-dioxane and agitated for 3 hours at 25 °C. The residue was concentrated and dissolved in DCM (5 ml). Acetyl-L-Leu-L-(OtBu-Glu)-L-Ala-L-Leu-L-Phe-COOH (130 mg, 0.19 mmol), HOBt (41.3 mg, 0.27 mmol), NMM (90 μl NEET) and EDC (54 mg, 0.28 mmol) were added sequentially and the mixture was stirred for 24 hours at 25 °C. The compound was concentrated, redissolved in dichloromethane (7.5 ml) and partitioned between H_2O (25 ml) and EtOAc (2 x 25ml). The organic layers were combined, dried over Na_2SO_4, concentrated and purified on a 40 ml flash chromatography column using 5% methanol in DCM. The compound was a pale yellow solid (155 mg, 0.278 mmol, 50% yield).

Ethyl 3-[Acetyl-L-Leu-L-Glu-L-Ala-L-Leu-L-Phe-L-Gln]-(E/Z)-propenoate (Compound F). **Compound E** (155 mg, 0.278 mmol) was added to a solution of TIS (140 μl) in TFA (1.85 ml) and stirred for 30 minutes. The compound was concentrated and washed with diethyl ether at -70 °C which provided a yellow crystalline solid (112 mg, 99%). The final product was produced in good yield with high purity and provided an expected mass of 816.77 Da (Figure 3.4). **Compounds A-F** were confirmed via electrospray mass spectrometry. Additionally, the Trt-Q-ethyl propenoate moiety and final product were dissolved in DMSO and their structures confirmed by 1D and 2D ^{1}H NMR (Figures 3.6, 3.7 and 3.9).

3.2.2 Mass Spectrometry

Electrospray mass-spectrometry was conducted on the reaction intermediates during the synthesis of the ethyl propenoate moiety and on the final acetyl-LEALFQ-ethyl propenoate peptidyl inhibitor to validate reaction products. ES-MS mass spectroscopy was done using a triple quad Micromass VG Quattro quadrupole MS unit. Samples were solubilized in 75% acetonitrile : 25% H_2O and loaded directly. Positive mode using 20 eV was used for parent ion detection. ES-MS of the final acetyl-LEAFLQ-ethyl propenoate inhibitor is shown in Figure 3.4.

MALDI-TOF mass spectrometry was done on an Applied Biosystems Voyager System

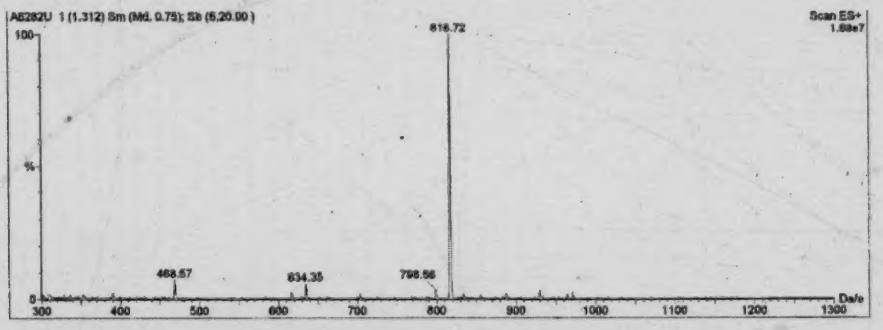

Figure 3.4: **ES-MS Spectrum of the Acetyl-LEALFQ-Ethylpropenoate Inhibitor**. The calculated mass of the final product is 816.76 Da. The ES-MS spectrum comfirms the presence of the expected parent ion at 816.72 Da.

6064 to confirm covalent modification of the $^{13}C/^{15}N$-labelled HRV14-3C protease with the synthesized acetyl-LEALFQ-ethyl propenoate inhibitor. Samples were applied to MALDI plates using a sinapinic acid matrix with C4-Ziptips. The acquisition mass range was set from 1,000 to 26,000 Da. A total of 29 and 78 shots were averaged for collection of the uninhibited and inhibited HRV14-3C protease spectra respectively (Figure 3.5). The accelerating voltage was set to 25,000 V, the grid voltage was 92% and the extraction delay time was 300 nsec.

3.2.3 NMR

Before continuing with the final peptide coupling, a sample of the tBoc-Q(Trt)-ethyl propenoate intermediate was dissolved in deuterated chloroform and analyzed by 1D ^1H NMR. The spectrum was collected on a Bruker AM-300 spectrometer with a spectral width of 2500 Hz (16,384 points). 200 transients were averaged and the sample was spun at 20 Hz. To characterize the final acetyl-LEALFQ-ethyl propenoate inhibitor, a 5 mM sample was prepared in deuterated DMSO. All spectra were collected at 25 °C and referenced to DMSO at 2.49 ppm. One dimensional ^1H (Figure 3.6) and a 2D ^1H-TOCSY (Figure 3.7) spectra were collected at 500 MHz using a Varian INOVA spectrometer.

Thirty-two transients were averaged for collecting these experiments. For the 1D spectrum, 11,470 points were collected over a 4200 Hz sweep width (8.4 ppm). The number of transients averaged was 256. Structural data was obtained from a 2D ^1H-NOESY spectum

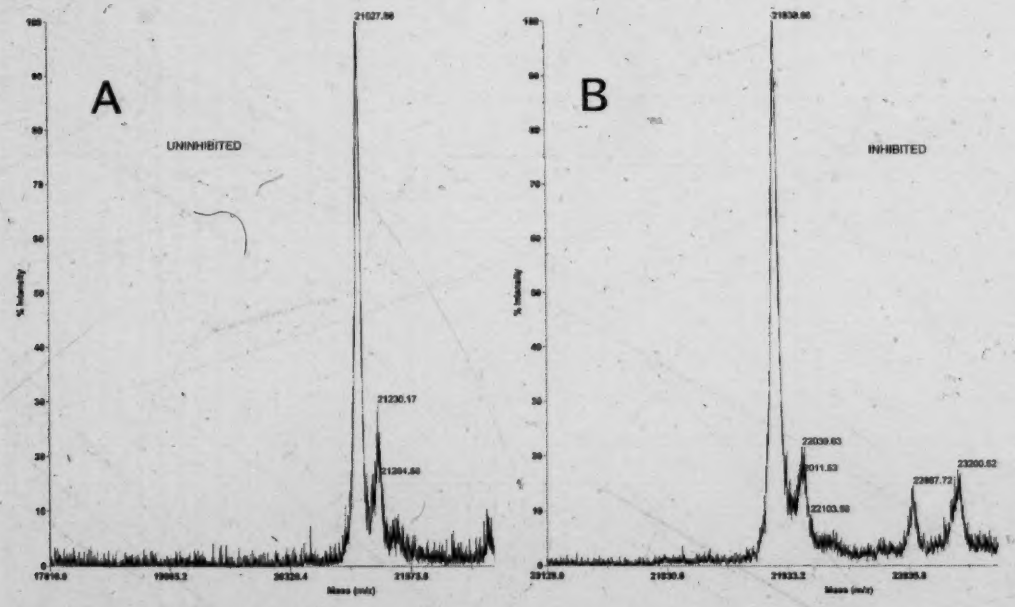

Figure 3.5: **MALDI-TOF Spectra of the [U-^{13}C/^{15}N]-HRV14-3C Samples.** Spectrum A is the reference for the [U-^{13}C/^{15}N]-HRV14-3C protease. Spectrum B is the [U-^{13}C/^{15}N]-HRV14-3C protease bound with the acetyl-LEALFQ-ethyl propionate inhibitor post dialysis. The mass difference between the [U-^{13}C/^{15}N]-labelled samples (~811 amu) confirmed the covalent modification of the protease and supports the enzymatic assay results (Figure 3.10).

collected at 500 MHz. 1024 and 512 points were collected over 6000 Hz in both the F2 and F1 dimensions and averaged over 32 transients. For the 2D ^{1}H-TOCSY spectrum, the sweep width was set to 5000 Hz in both dimensions. The number of points collected in F1 and F2 dimensions was 256 and 1024 respectively. A ^{13}C-HSQC spectrum (Figure 3.9) was also collected. For this experiment, sweep widths of 4000 and 27,000 Hz (8 and 216 ppm) were collected using 1024 and 256 points in the ^{1}H and ^{13}C dimensions respectively.

3.3 Results and Discussion

3.3.1 Inhibitor Synthesis

To synthesize the HRV14-3C inhibitor, the methods described by Dragaovich *et al.* [6] were adapted to generate the tBoc protected Gln(Trt)-ethyl propenoate group. Modifica-

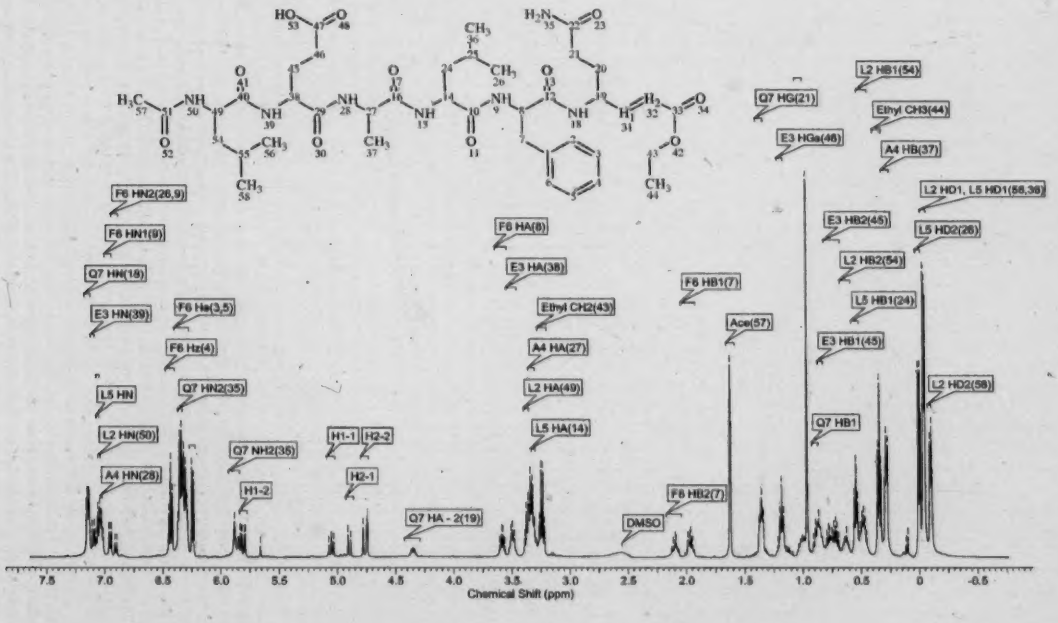

Figure 3.6: **Chemical Structure and 1D ^{1}H NMR Spectrum of the Acetyl-LEALFQ-Ethyl Propenoate Inhibitor in DMSO**. The connectivity was resolved with the 2D ^{1}H -TOCSY spectrum (τ_m 60 ms) shown in Figure 3.7. Labels are shown above the peaks along with proton labels in parenthesis that correspond to the proton labels mapped onto the structure. The chemical shift assignments are listed in table D.2.

tions to this protocol were primarily dictated by reagent availability. The (acetyl-LEALF-COO^-) peptide was selected based on the natural peptide substrates used in the development of the commercially available HRV14-3C pNA assay [5]. It was prepared *via* Fmoc solid phase peptide chemistry on a 2-chlorotrityl resin, which allowed cleavage of the peptide using weak acid while retaining the necessary tBoc and Trt protecting groups for the solution phase coupling to the Gln(Trt)-ethyl propenoate group.

The solution phase, peptide-coupling reaction was about 50% efficient. Following this reaction, the remaining protecting groups were removed with TFA to yield 102 mg (0.125 mmol) of final pure product. Complications during the final solution coupling were encountered. ES-MS spectra of the reaction mixture showed failure in tBoc deprotection of the Gln(Trt)-ethyl propenoate group, which was resolved by increasing the acid concentration from 2 M to 4 M HCl [7]. The 1D ^{1}H NMR spectroscopy on the purified tBoc-[Q(Trt)]-ethyl propenoate intermediate was performed to assess the authenticity of the product before con-

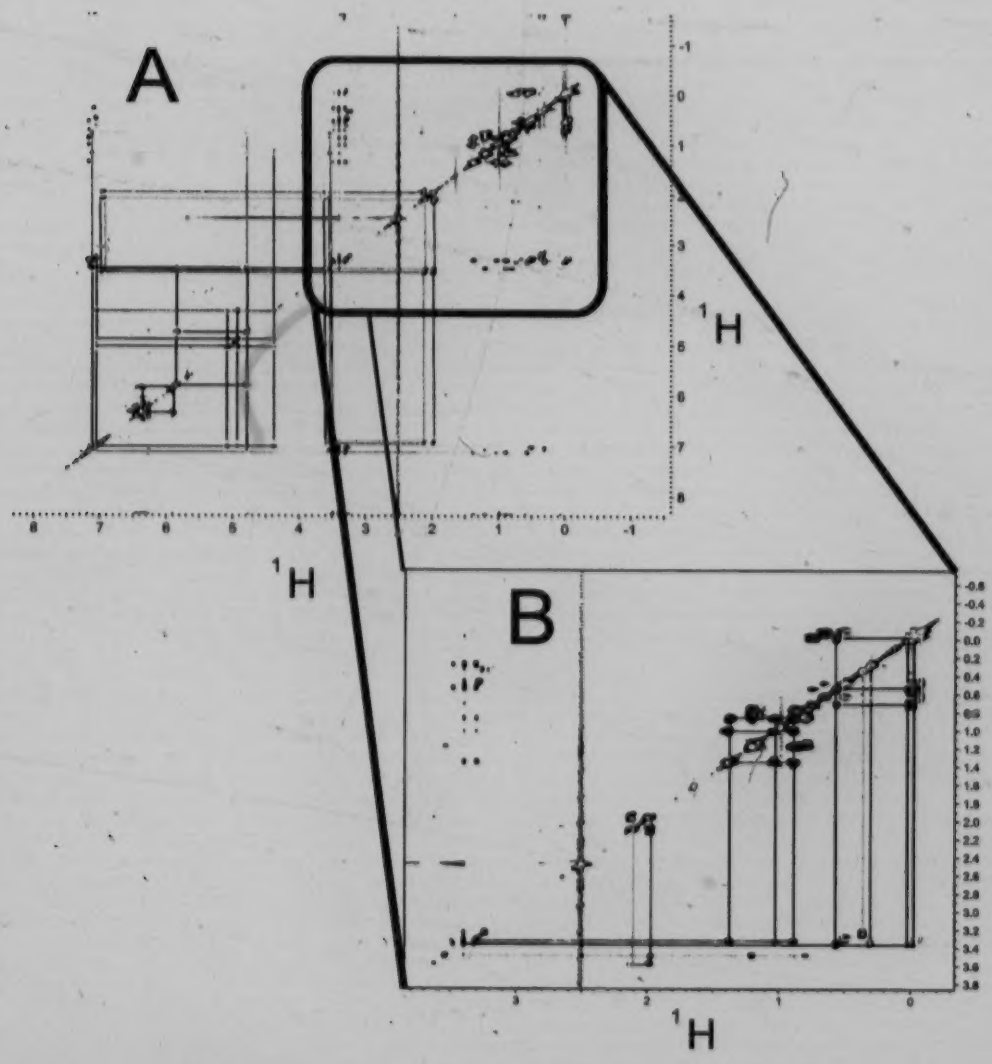

Figure 3.7: **^{1}H-TOCSY of the Acetyl-LEALFQ-ethyl Propenoate Inhibitor in DMSO**. The full ^{1}H-TOCSY spectrum is shown in spectum A. The aliphatic region is expanded in spectrum B. The corresponding chemical shift assignments are listed in table D.2. Lines connecting the spin systems have to following color code: Spectrum A. Black - trans alkene; Light Purple - cis alkene; Purple - P_1 side-chain amide; Light blue - P_1 Gln; Green and Orange - P_2 Phe. Spectrum B. Light blue - P_1 Gln; Orange and Green - P_2 *D/L* Phe; Black - P_3 Leu; Purple - P_4 Ala; Red - P_5 Glu; Yellow - P_6 Leu.

tinuing onto the final solution phase coupling to the peptide. From coupling constant data, it was evident that isomerization of the double bond had occurred (~60% trans and 40% cis based on ^{1}H NMR peak integration). This was later confirmed with subsequent one and two dimensional ^{1}H spectra collected on the final products (Figures 3.6 and 3.7). These different conformations influenced not only the $^1H_\alpha$ shift of the P_1 Gln, but the $^1H_\alpha$ shift of the P_2 Phe. This observation was confirmed from the 2D ^{1}H-NOESY spectrum that indicated these chemical shift differences resulted from the proximity of the ester relative to the P_1 and P_2 residues. Table D.1 lists the ^{13}C chemical shifts obtained from the ^{13}C -HSQC spectrum shown in Figure 3.9. The chemical shift differences that result from the two diastereoisomers are clearly visible in this spectrum. The spectrum shows two very different $^1H_\alpha$ and $^{13}C_\alpha$ peaks for the P_1 glutamine result from the proximity of the atoms relative to the electron withdrawing ester group. The *i-1* phenylalanine is sufficiently close to also have its $^1H_\alpha$ peaks affected. The 1D ^{1}H spectrum (Figure 3.6) was used to quantify the cis/trans isomerization and measure vincinal coupling constants. Figure 3.8 depicts the relative orientation of the ethyl propionoate group in relation to the P_1 $^1H_\alpha$ proton for both the cis and trans isomers. The proximity of the P_1 $^1H_\alpha$ to the ethyl propionoate ester group accounts for the downfield chemical shift observed in the trans isomer.

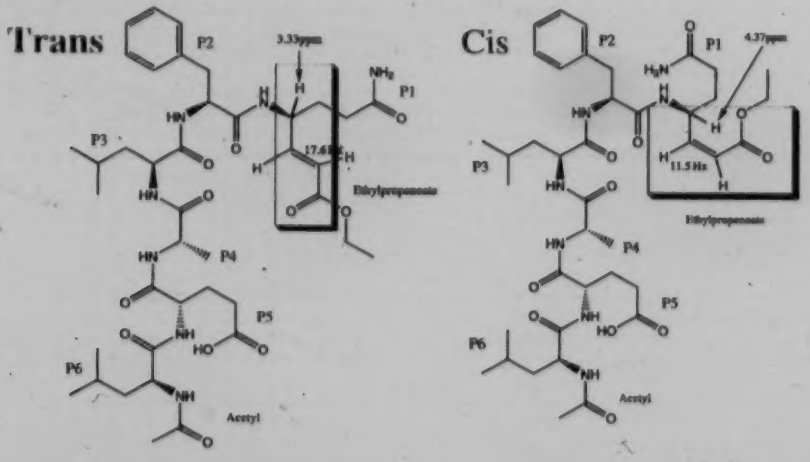

Figure 3.8: **Cis/Trans Diastereoisomers of the Ethyl Propenoate Group**. Differences in vicinal coupling is observed between the trans and cis isomers of the ethyl propenoate group (~6 Hz). The cis isomer positions the electron withdrawing ester group closer to the P_1 and P_2 $^1H_\alpha$ protons and likely accounts for the 1.04 ppm chemical shift difference between the two isomers.

Previously published protocols of this synthesis [6] did not report the production of stereo isomers (reaction **C**: Figure 3.3). However, the E/Z diastereomer of the ethylpropenoate alkene (~60 trans and ~40 cis) was within expected ratios due to the use of a stabilized ylide [8] and the substituted groups on the α-branched phosphonate [9], which was employed in the Horner-Wadsworth-Emmons reaction [10]. To work around concerns regarding the E/Z isomers a 6-fold molar excess of inhibitor was incubated with the enzyme during covalent modification. This assured a 3.6-fold excess of inhibitor as only the E isomer is suspected to be biologically active because the Z stereoisomer might encounter steric hindrance in the S′ substrate pockets. This ratio is similar to the ratio (3-fold) used to inhibit the HRV2-3C protease with AG7088 [1].

3.3.2 Inhibitor Solubility and Enzyme Activity Assays

The p-nitroaniline assay [5] described in section 2.2.2 was used to screen various co-solvent mixtures used to dissolve the acetyl-LEALFQ-ethyl propenoate inhibitor and to check for enzyme inhibition following the incubation with the 6-fold molar excess of the E/Z-inhibitor. An inhibitor peptide length of six residues was required to assess possible interactions within the S_5 and S_6 substrate pockets. However, increasing the inhibitor length a further two residues and including a lipophilic amino acid in the P_6 position presented solubility issues. Solubility problems were also encountered in the study that led to the development of the commercially available p-nitroaniline peptide [5]. This work utilized a modified 2C/3A cleavage site analogue (EALFQ vs. ETLFQ) to improve solubility. Despite adopting this substitution in our inhibitor design, solubility problems persisted (predicted LogP of ~2.1). This problem was overcome by testing various co-solvent mixtures for the inhibitor and enzyme-inhibitor reactions. Previously, a 2% v/v DMSO/H_2O was used as a co-solvent for the incubation reaction used to inactivate the HRV2-3C protease with AG7088 [1]. However, Wang and Jonhson [11], reported reduced enzyme activity for the HRV14-3C protease when using DMSO co-solvent mixtures, presumably from the oxidation of the active site cysteine residue (Cys^{146}). Our studies confirmed their results with a 30% reduction of HRV14-3C protease activity in 2.8% v/v DMSO/H_2O co-solvent blank compared with water alone (Figure 3.10). Their work subsequently explored low concentration methanol-water and isopropyl alcohol-water co-solvent systems, which presented no significant reduction

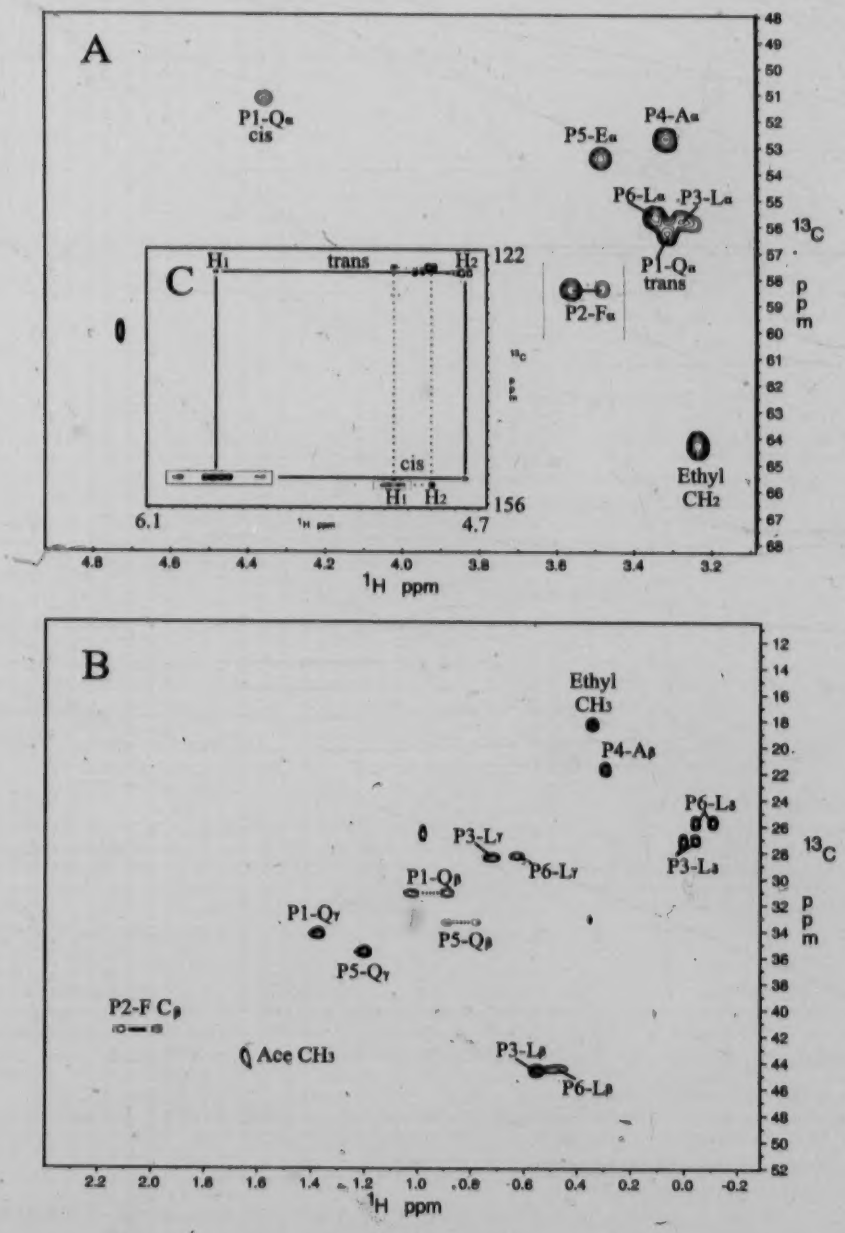

Figure 3.9: **^{13}C-HSQC of the Acetyl-LEALFQ-ethyl propenoate Inhibitor in DMSO**. The $^{13}C_\alpha$ region is shown in spectrum A. The inserted spectrum C shows the downfield shifted alkene region. Spectra A and B shows different portions of the aliphatic region. The corresponding chemical shift assignments are listed in Table D.1. Lines connect geminal protons. Conformational differences between the cis and trans isomers presented $^1H_\alpha$ and $^{13}C_\alpha$ chemical shift differences for the P_1 glutamine and $^1H_\alpha$ chemical shift differences for the P_2 phenylalanine (boxed). The labelling scheme corresponds to: substrate position-amino acid-atom.

of protease activity. We also found that the protease activity was not altered in water with ethanol concentrations over 2.6% (well above the concentration of the final enzyme-inhibitor reaction) and marginally decreased (~10%) in 10% v/v ethanol/water. These results are consistent with Wang and Johnson's findings [11]. Consequently, ethanol/water was chosen for inhibitor-enzyme reactions with the inhibitor initially being disolved in 43% v/v ethanol/water at a concentration of 2.5 mM. This allowed the addition of a 6-fold molar excess of inhibitor for the colorimetric assays with an ethanol concentration that did not exceed 2%. The enzyme inhibition assay was done by addition of a 6-fold molar excess of inhibitor after 1 minute of reaction equilibration. Complete inhibition was evident from the lack of increased absorbance (UV_{405nm}) following the addition of the inhibitor (final ethanol concentration 1.7%). Covalent modification of the HRV14-3C enzyme was further confirmed by repeating the assay on the [U-$^{13}C/^{15}N$]-HRV14-3C sample following dialysis of the enzyme against 20 mM KH_2PO_4 buffer (0.5 mM EDTA, pH 6.5) whereby no detectable absorbance of UV_{405nm} was observed.

Final confirmation of enzyme modification was made with MALDI-TOF mass spectrometry on both the inhibited [U-$^{13}C/^{15}N$]-HRV14-3C NMR sample and a sample of the same apo, labelled protease. Comparison of the two samples confirmed covalent modification with the mass of the bound protein-inhibitor complex being 21839 Da and the mass of the free form being 21028 Da (Figure 3.5). The difference between these numbers (811 amu) is within 0.03% of the expected mass difference or mass of the inhibitor (817 amu) for the labelled protein sample. Given the fact that these numbers compare two 'uniformly' labelled protein samples and these respective numbers were obtained by averaging a series of scans, the relatively small difference between the expected and observed difference is not significant.

3.4 Conclusion

The main objective for this component of the project was to synthesize a biologically active inhibitor for the HRV14-3C protease. This objective was met with the confirmation of inhibitor activity and covalent attachment to the HRV14-3C protease (Figures 3.10 and 3.5). Despite the problems with diastereomer formation during synthesis (Figures 3.7 and

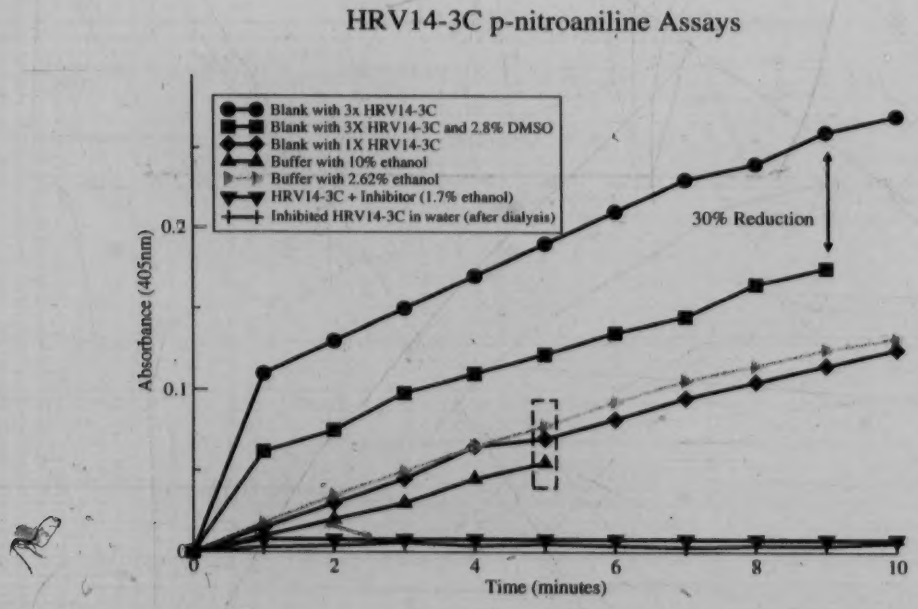

Figure 3.10: **HRV14-3C p-Nitroaniline Colorimetric Assays**. These assays were done at 25 °C in KH_2PO_4 buffer/co-solvent mixtures at pH 7.0 in the absence and presence of inhibitor (62.5 μM EALFQ-p-nitroaniline) to confirm activity and inactivation of the HRV14-3C protease (1x: 3.45 μM for ethanol co-solvent analysis; 3x: 10.35 μM for DMSO analysis). The • 3x HRV14-3C blank and ■ 2.8% DMSO co-solent reactions show oxidation of the enzyme. Boxed: ♦ 1x HRV14-3C blank, ▲ 10% ethanol co-solvent and ► 2.62% ethanol co-solvent show enzyme stability in buffer with >2.5% ethanol. The ▼ 6-fold molar excess of inhibitor added post 1 minute reaction (final ethanol concentration 1.7%) and + Blank with inhibited enzyme reactions confirm inactivation and covalent modification.

3.9 and Table D.2), complete inactivation of the protein was afforded by incubation with a 6-fold molar excess of inhibitor to enzyme. The excessive amount of inhibitor required for this reaction was prepared in a 2.5 mM stock solution in a 43% ethanol/water co-solvent. The resulting final ethanol concentration in the incubation reaction was 1.7%, which was well below levels shown to impair enzyme activity (Figure 3.10).

The assessment of post-scissile bond interactions was forgone in favor of the covalent modifying ethyl propenoate moiety, which proved to be successful for the study of the HRV2-3C protease [1]. This strategy allowed the addition of a peptide large enough to interact with the S_5 and S_6 pockets and would afford interaction with Asn^{164}. This asparagine residue is

of particular interest because it is involved in a previously characterized deamidation event for the HRV14-3C protease [12].

Initial concerns about enzyme stability were put to rest as the concentrated inhibited HRV14-3C protease remained in solution following the collection of preliminary 3D heteronuclear NMR spectra (HNCA and HNCACB). Furthermore, no changes in the ^{15}N-HSQC spectra of the inhibited enzyme were apparent even 1 year after the sample was prepared. There was neither shifting of amide peaks nor emerging spurious signals in the random coil region of the spectrum. The only difference between the apo and inhibited HRV14-3C ^{15}N-HSQC spectra appeared for residues near the proteolytic site. Furthermore, an increased number of amide signals were present in the newly acquired ^{15}N-HSQC data for the inhibited enzyme that suggested stabilization of the proteolytic site might have taken place and that a detailed structural analysis of the proteolytic pharmacophore might be possible. These data are presented in the following chapters that discuss the complete chemical shift assignments and structure calculations for both the apo and inhibited forms of the HRV14-3C enzyme.

Bibliography

[1] D. A. Matthews, P. A. Dragovich, S. A. Webber, S. A. Fuhrman, A. K. Patick, L. S. Zalman, T. F. Hendrickson, R. A. Love, T. J. Prins, J. T. Marakovits, R. Zhou, J. Tikhe., C. E. Ford, J. W. Meador III, R. A. Ferre, E. L. Brown, S. L. Binford, M. A. Brothers, D. M. DeLisle, and S. T. Worland. Structure-Assisted Design of Mechanism-Based Irreversible Inhibitors of Human Rhinovirus 3C Protease with Potent Antiviral Activity Against Multiple Rhinovirus Serotypes. *Proc Natl Acad Sci U S A*, 96:11000–11007, 1999.

[2] S. C. Mosimann, M. M. Cherney, S. Sia, S. Plotch, and M. N. G. James. Refined X-ray Crystallographic Structure of the Poliovirus 3C Gene Product. *J Mol Biol*, 273:1032–1047, 1997.

[3] E. M. Bergmann, M. M. Cherney, J. McKendrick, J. C. Vederas, and M. N. G. James. Crystal Structure of an Inhibitor Complex of the 3C Proteinase from Hepatitis A Virus (HAV) and Implications for the Polyprotein Processing in HAV. *Virology*, 265:153–163, 1999.

[4] J. Yin, E. M. Bergmann, M. M. Cherney, S. M. Lall, R. P. Jain, J. C. Vederas, and M. N. G. James. Dual Modes of Modification of Hepatitis A Virus 3C Protease by a Serine-derived Beta-Lactone: selective crystallization and Formation of a Functional Catalytic Triad in the Active Site. *J Mol Biol*, 354:854–871, 2005.

[5] Q. M. Wang, R. B. Johnson, G. A. Cox, E. C. Villarreal, and R. J. Loncharich. A Continuous Colorimetric Assay for Rhinovirus-14 3C Protease Using Peptide p-Nitroanilides as Substrates. *Anal Biochem*, 252:238–245, 1997.

[6] P. S. Dragovich, S. E. Webber, R. E. Babine, S. A. Fuhrman, A. K. Patick, D. A. Matthews, C. A. Lee, S. H. Reich, T. J. Prins, J. T. Marakovits, E. S. Littlefield, R. Zhou, J. Tikhe, C. E. Ford, M. B. Wallace, J. W. Meador III, R. A. Ferre, E. L. Brown, S. L. Binford, J. E. V. Harr, D. M. DeLisle, and S. T. Worland. Structure-Based Design, Synthesis, and Biological Evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. 1. Michael Acceptor Structure-Activity Studies. *J Med Chem*, 41:2806–2818, 1998.

[7] P. S. Dragovich, R. Zhou, D. J. Skalitzky, S. A. Fuhrman, A. K. Patick, C. E. Ford, J. W. Meador III, and S. T. Worland. Solid-phase Synthesis of Irtreversible Human Rhinovirus 3C Protease Inhibitors. Part I: Optimization of Tripeptides Incorporating N-terminal Amides. *Bioorg and Med Chem*, 7:589–598, 1999.

[8] J. Clayden, N. Greeves, S. Warren, and P. Wothers. *Organic Chemistry.* Oxford University Press, Oxford, NY, USA, 2001.

[9] H. Nagaoka and Y. Kishi. Further Synthetic Studies on Rifamycin S′. *Tetrahedron*, 37:3873–3888, 1981.

[10] W. S. Wadsworth and W. D. Emmons. The Utility of Phosphonate Carbanions in Olefin Synthesis. *J Am Chem Soc*, 83:1733–1738, 1961.

[11] Q. M. Wang and R. B. Johnson. Activation of Human Rhinovirus-14 3C Protease. *Virology*, 280:80–86, 2001.

[12] G. A. Cox, R. B. Johnson, J. A. Cook, M. Wakulchik, M. G. Johnson, E. C. Villarreal, and Q. M. Wang. Identification and Characterization of Human Rhinovirus-14 3C Protease Deamidation Isoform. *J Biol Chem*, 274:13211–13216, 1999.

Chapter 4

Chemical Shift Assignment and Solution Structure Calculation of the Inhibited HRV14-3C Protease

4.1 Introduction

As mentioned in the previous chapter, initial experimental work conducted on the inhibited HRV14-3C protease revealed some interesting results. The protein exhibited improved stability upon inhibition as hypothesized and the extent of this stability enhancement was surprising. In fact, the original sample initially prepared for the collection of NMR experiments to determine backbone chemical shifts remains in solution to date (4 years after the initial sample preparation) and no change in ^{15}N-HSQC spectra is evident indicating that this sample remains free from degradation. All the experiments listed in Table 4.1 were performed on the same inhibited [U-^{13}C/^{15}N]-HRV14-3C protease sample. The quality of NMR data for the inhibited form is substantially better compared with the data collected on the apo HRV14-3C samples. The higher quality of NMR spectra and excellent solution state stability of the inhibited HRV14-3C protease enabled the the structure to be completed and the pharmacophore of the enzyme to be analyzed in detail. This chapter outlines the experiments and methodology used to complete the inhibited HRV14-3C structure and extensively analyzes the pharmacophore of the proteolytic site. Finally, the peptide-inhibitor substrate bound to HRV14-3C is compared to the peptide-mimic inhibitor, AG7088, bound to the homologous HRV2-3C enzyme and sub-genus differences are discussed.

4.2 Materials and Methods

4.2.1 Sample Preparation

NMR samples were prepared using the methods outlined in *Appendix B* with the recipes listed in *Appendix A*. [U-^{15}N]-labelled and [U-^{13}C/^{15}N]-labelled samples were prepared by substituting ^{15}N-labelled NH_4Cl and/or ^{13}C-labelled glucose for the unlabelled ingredients listed in the MM recipe listed in *Appendix A*. NMR samples were prepared by dialyzing the purified protease into 20 mM KH_2PO_4 buffer (pH 6.5, 0.5 mM EDTA, 15 mM DTT). Acetyl-LEALFQ-ethyl propionate inhibited HRV14-3C samples were concentrated to ~0.75 mM. DSS (0.1 mM) was added for internal referencing. D_2O (10%) was added to maintain the spectrum lock. Samples were filtered through 22 μm epindorph filters and placed in either 5 mm thin walled WILMAD© NMR tubes and sealed under argon or 3mm SHIGEMI© NMR tubes. Subsequent to collecting the majority of experiments (backbone, side-chain and NOESY), the sample was exchanged into 20 mM KH_2PO_4 buffer prepared in 99.9% D_2O (pD 6.9, 0.5 mM EDTA, 15 mM DTT, 0.1 mM DSS, final D_2O concentration: 99.6%) *via* successive volumetric dilutions and ultracentrifugation to obtain ^{1}H/^{2}H data and collect an additional ^{13}C -NOESY-HSQC spectrum. All the experiments collected on the inhibited HRV14-3C protease are listed in Table 4.1.

4.2.2 NMR Experiments

All NMR experiments collected on the inhibited HRV14-3C enzyme were conducted on a Varian INOVA 500 MHz spectrometer equipped with either a room temperature 5 mm triple-resonance z-axis PFG probe or a 5 mm triple-resonance z-axis PFG cold probe. All experiments were conducted using either Varian ProteinPack or BioPack pulse sequences (VNMR v3.1c) with the exception of the 2D ^{13}C/^{15}N-F1/F2-Filtered TOCSY pulse sequence, which was supplied by Dr. Leo Spyracopoulos (University of Alberta). All spectra were collected at 25 °C, referenced indirectly to DSS [1], processed with NMRPIPE [2] and analyzed with NMRVIEW [3]. The experiments and spectral parameters used for chemical shift and restraint assignments for the inhibited HRV14-3C protease are shown in in Table 4.1.

Table 4.1: **NMR Experiments Conducted on the Inhibited HRV14-3C Protease**

Experiment	Nucleus t1	Nucleus t2	Nucleus t3	Number of Points t1	Number of Points t2	Number of Points t3	Spectral Width (Hz) t1	Spectral Width (Hz) t2	Spectral Width (Hz) t3	Transients
Backbone Assignments										
^{15}N-HSQC*	^{15}N	^{1}H		256	1024		2200	8000		32
HNCO	^{13}C	^{15}N	^{1}H	64	32	1024	3018	2000	6000	16
HNHA†	^{1}H	^{15}N	^{1}H	64	32	1024	6000	2000	6000	16
HNCA	^{13}C	^{15}N	^{1}H	64	32	1024	3770	2000	6000	16
HNCACB	^{13}C	^{15}N	^{1}H	96	32	1026	10056	2200	6000	32
CBCA(CO)NNH	^{13}C	^{15}N	^{1}H	48	32	1024	10056	2200	6000	32
Sidechain Assignments										
C(CO)NNH*	^{13}C	^{15}N	^{1}H	128	32	768	10054	2200	6000	16
H(C CO)NNH*	^{1}H	^{15}N	^{1}H	64	32	1024	6000	2000	6000	32
HCCH-TOCSY†	^{1}H	^{13}C	^{1}H	64	32	1024	6000	2000	6000	32
CCH-TOCSY†	^{13}C	^{13}C	^{1}H	64	32	1024	8000	10000	6000	32
2D ^{1}H-^{1}H TOCSY†‡	^{1}H	^{1}H		512	1024		6000	6000		16
^{13}C HSQC (35 ppm§)	^{13}C	^{1}H		256	1024		12568	6000		48
^{13}C HSQC (125 ppm§)	^{13}C	^{1}H		64	512		7542	6000		16
Restraint Assignments										
^{15}N-NOESY-HSQC (τ_m 75 ms¶)	^{1}H	^{15}N	^{1}H	64	32	1024	6000	2200	6000	32
^{13}C-NOESY-HSQC (35ppm§, τ_m 100 ms¶)†‡	^{1}H	^{13}C	^{1}H	64	64	1024	6000	10060	6000	32
^{13}C-NOESY-HSQC (125ppm§, τ_m 100ms¶)‡	^{1}H	^{13}C	^{1}H	48	32	1024	6000	1000	6000	32
D_2O exchange ^{15}N HSQCs‡	^{15}N	^{1}H		128	1024		2200	6000		32
Inhibitor Specific Assignments										
2D ^{13}C/^{15}N-Filtered TOCSY† (τ_m 50 ms¶)	^{1}H	^{1}H		512	1024		6000	6000		128
2D ^{13}C/^{15}N-Filtered NOESY† (τ_m 100 ms¶)	^{1}H	^{1}H		128	2048		6000	6000		128
^{13}C-Filt./Edit NOESY-HSQC†‡ (τ_m 100 ms¶)	^{1}H	^{13}C	^{1}H	64	32	1024	6000	6090	6000	32

*Collected in duplicate with room-temperature and cold probes
†Conducted with the cold probe
‡Conducted on the 99.6% D_2O sample
§Carrier frequency
¶Mixing time

4.2.3 NMR Chemical Shift and NOE Restraint Assignment

The sequential assignment process was aided with the availability of previously published backbone chemical shifts for the apo HRV14-3C protease [4]. The backbone chemical shift assignment data for the inhibitor-bound form of HRV14-3C was obtained from HNCA [5], HNCACB [6], CBCA(CO)NNH [7], HNCO [8] and HNHA [9] experiments. Side chain assignments were completed from HC(CO)NNH [10], C(CO)NNH [10], HCCH-TOCSY [11, 12], and CCH-TOCSY [11] spectra. Aromatic assignments were obtained from an aromatic ^{13}C-HSQC spectrum [13] with a carrier frequency of 125 ppm and a 2D ^{1}H watergate-TOCSY [14] collected on the [U-^{13}C/^{15}N]-HRV14-3C sample in 99.6% D_2O. Aromatic proton assignments were confirmed with $^{1}H_{\beta}$ to $^{1}H_{\delta}$ NOE cross-peak identification using the ^{13}C-edited NOESY spectrum collected on the [U-^{13}C/^{15}N]-labelled HRV14-3C sample in 99.6% D_2O. Bound inhibitor chemical shift assignments were obtained from the

2D $^{13}C/^{15}N$-F1/F2-Filtered TOCSY.

4.3 Results

4.3.1 Inhibited HRV14-3C Chemical Shift Assignments

NMR experiments that correlated the $^{1}H_{N}$ peaks (identified from the ^{15}N-HSQC spectrum shown in Figure 4.1) to intra-residue (*i*) and inter-residue (*i-1*) backbone and side-chain chemical shift data were used to complete the chemical shift assignments. Nearly all the expected amide signals were obtained from the ^{15}N-HSQC spectrum indicating the sample had improved stability and possibly altered dynamics compared with the apo form. The most notable difference from comparing the ^{15}N-HSQC spectra obtained from the inhibited enzyme and the apo enzyme (Figure 5.1) were the lack of spurious degradation signals (ν_N ~122 ppm, ν_{HN} ~8.25 ppm). The spectra shown in Figure 4.1B was obtained 1 year following the initial sample preparation. The principle NMR experiments that governed the assignment process were the HNCA and HNCACB, C(CO)NNH and H(C CO)NNH experiments. The HNCA and HNCACB experiments, which correlate both preceding (*i-1*) and intraresidue (*i*) $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts were used to sequentially link adjacent residues. To provide the remaining spin system information and resolve assignment ambiguities, the C(CO)NNH and H(CCO)NNH experiments were used. Examples of these spectra for the inhibited HRV14-3C protease are shown in Figures 4.2 and 4.3. These strip plots shown are for a region of the protein that presented assignment difficulties for the apo enzyme. Although the data is weaker in intensity compared with other areas of the protein (Figure 4.4), the data was attainable none the less. Information not readily resolvable from these experiments (ie. $^{13}C_{\gamma}$ and $^{13}C_{\delta}$ for Leu^{123} and the $^{13}C_{\delta}$ for Pro^{117}) were pieced together from CCH-TOCSY and HCCH-TOCSY experiments, which correlated side-chain ^{13}C-^{13}C and ^{1}H-^{1}H shifts (Figure 4.5). Both aliphatic and aromatic ^{13}C-HSQC spectra were also collected and used to assign and confirm some ^{1}H and ^{13}C peak assignments (ie. proline $^{13}C_{\delta}$ and glycine $^{13}C_{\alpha}$).

A 2D ^{1}H-TOCSY collected on the [U-$^{13}C/^{15}N$]-HRV14-3C sample in D_2O was used to connect aromatic spin systems, whose ^{1}H and ^{13}C chemical shifts were initially identified in the ^{13}C-edited HSQC experiment (carrier 125 ppm). This spectrum also resolved all side-

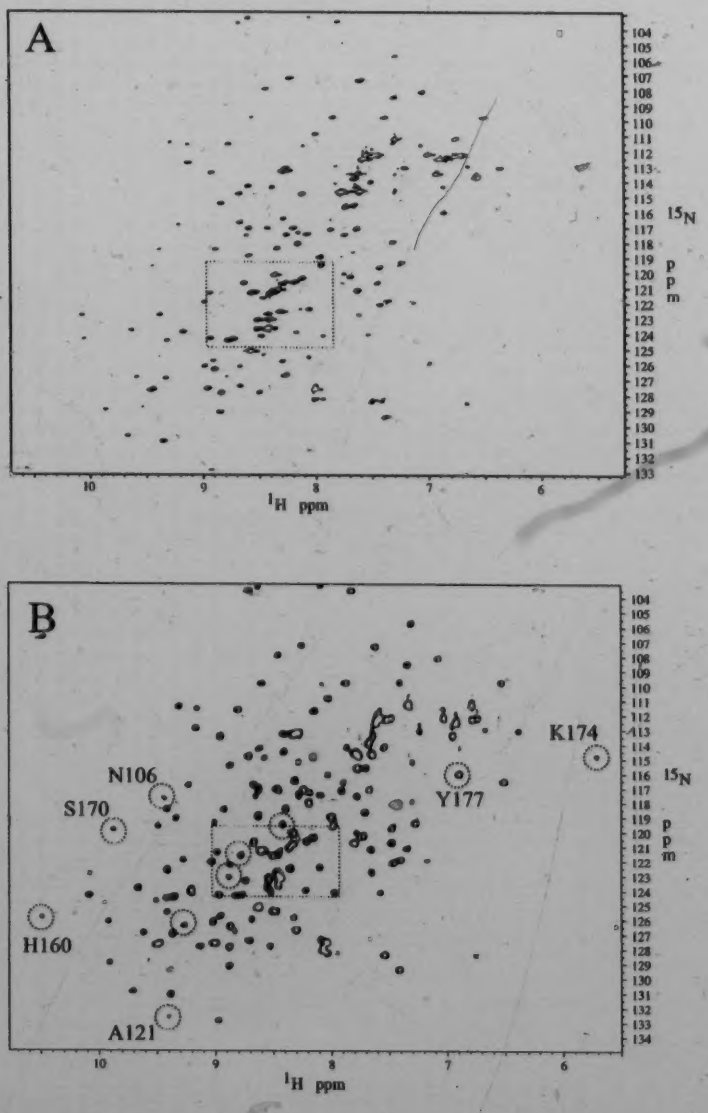

Figure 4.1: **Inhibited HRV14-3C ^{15}N-HSQC spectra. Spectrum A** was collected using a Varian 500 MHz INOVA spectrometer fitted with a room temperature HCN triple resonance probe. **Spectrum B** was collected 1 year later on the same sample, using the same instrument fitted with a cold probe. The lack of degradation peaks observed in the random coil region of this spectrum (ν_N ~122; ν_{HN} ~8.25 ppm) following this time period confirmed the sample stability. This region is boxed. In addition, a number of extra peaks are visible due to the increased sensitivity afforded with the cold probe. Some of these peaks have been highlighted with red circles and some include important active site residues such as His^{160}. These circled peaks were also visible in Spectrum A, however, not at the contour level shown.

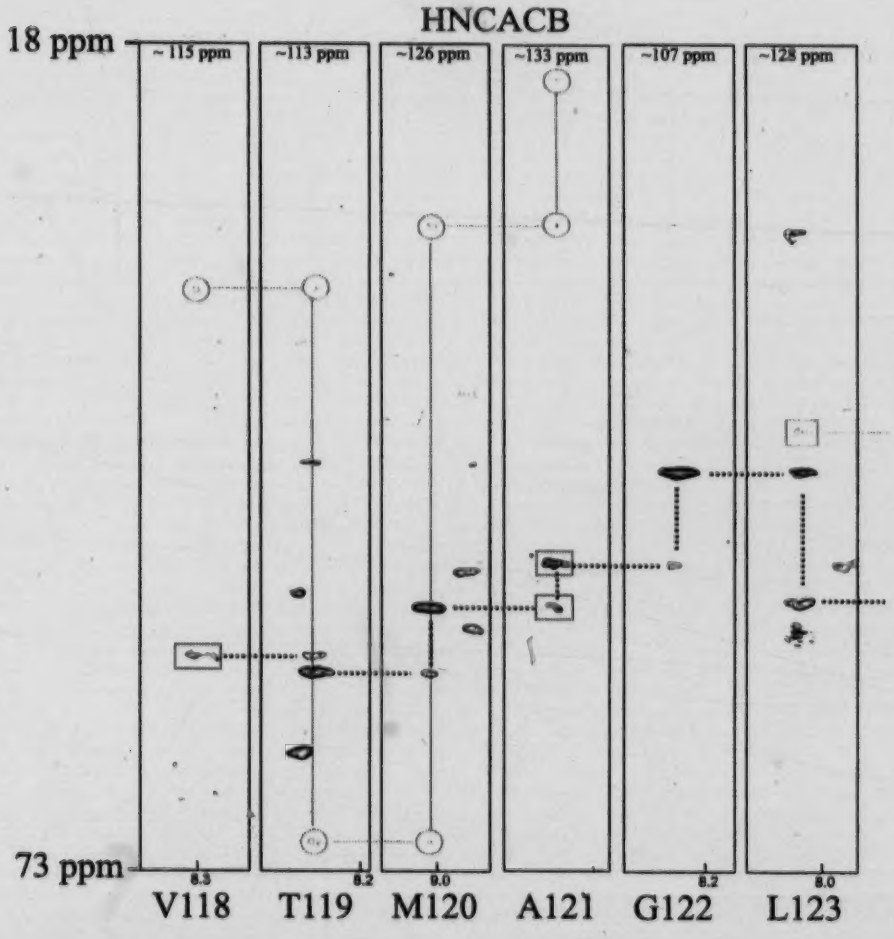

Figure 4.2: **HNCACB Strip Plots from the Inhibited HRV14-3C**. A walk-through for a portion of the HNCACB spectra that presented weaker peaks. This spectra was collected using a 500 MHz Varian INOVA spectrometer fitted with a room temperature HCN probe. $^{13}C_{\alpha}$ peaks are colored black and $^{13}C_{\beta}$ peaks are colored red. Broken lines join inter-residue *i* with *i-1* $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ peaks. Nuclei that present exchange broadening are boxed. The C(CO)NNH experiment (Figure 4.3) was used to piece together the preceding residue chemical shifts that were not observable in this experiment. The corresponding ^{15}N chemical shifts for each strip are listed at the top of each strip plot.

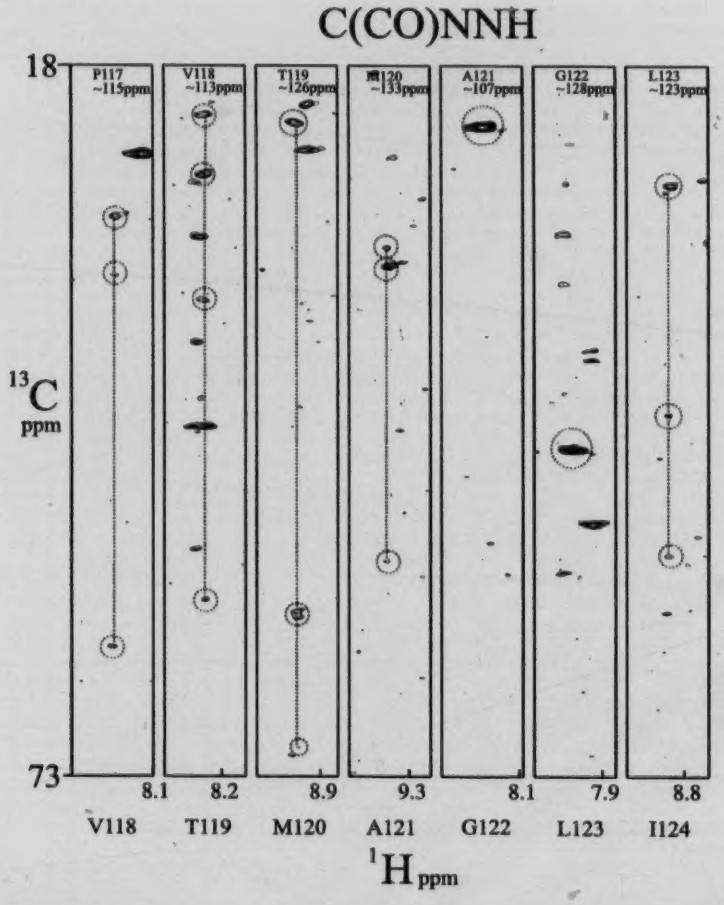

Figure 4.3: **C(CO)NNH Strip Plots from the Inhibited HRV14-3C**. The C(CO)NNH experiment, which correlates the *i-1* ^{13}C chemical shifts to the *i* $^{1}H_N$ chemical shifts, was used to fill in the remaining spin system information and resolve ambiguities that resulted from poor signal-to-noise in regions of the HNCACB experiment (Figure 4.2). Labels for the preceding residue that the strip plots provide chemical shift information for are inserted at the top of each strip plot. The $^{13}C_\delta$ peak for Pro^{117} is not visible at this contour level. The spectrum was collected using a 500 MHz Varian INOVA spectrometer fitted with a room temperature HCN probe.

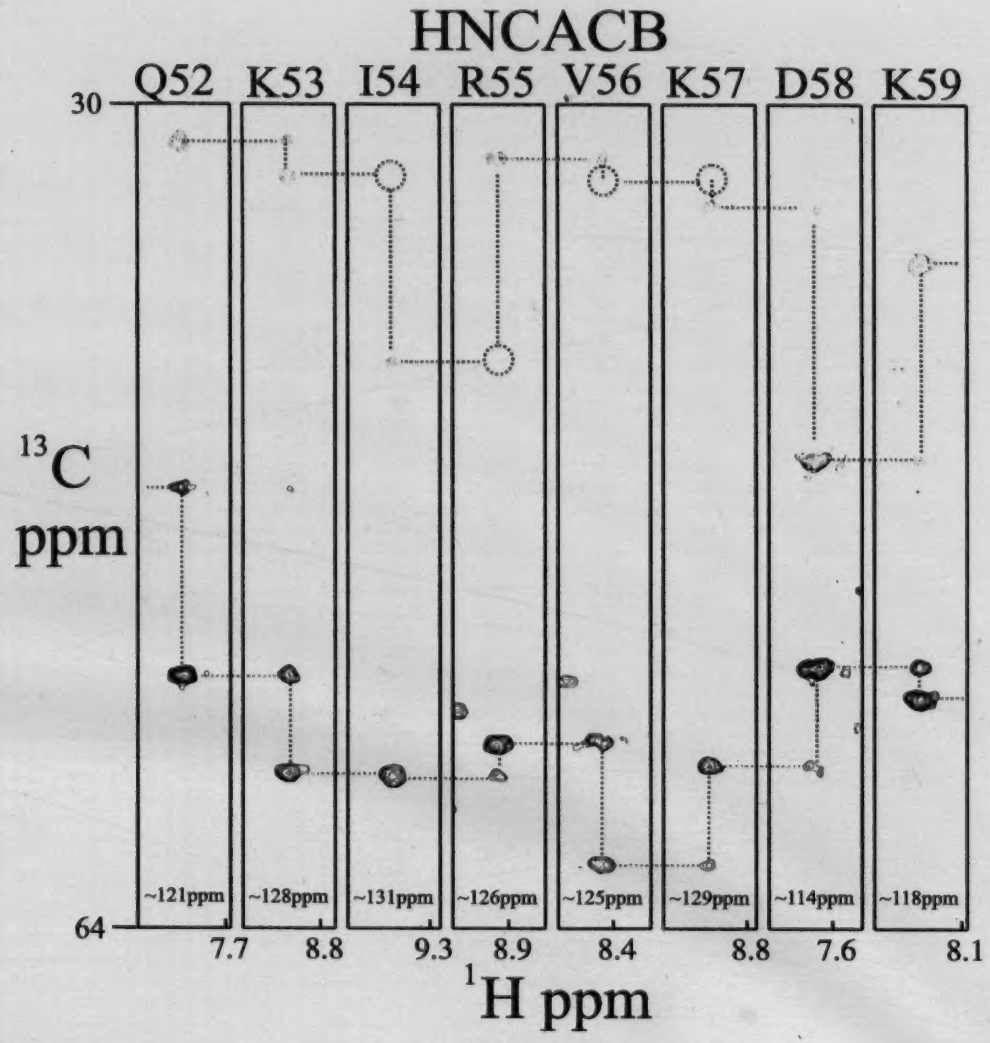

Figure 4.4: **HNCACB Strip Plots for Region QKIRVKDK$^{52-59}$ of the Inhibited HRV14-3C Protease**. A walkthough for a portion of the HNCACB spectra that presented improved signal-to-noise compare with the region shown in Figure 4.2. $^{13}C_{\alpha}$ peaks are colored black and $^{13}C_{\beta}$ peaks are colored red. Broken lines connect *i* and *i-1* peaks. The *i-1* $^{13}C_{\beta}$ chemical shifts not visible at this contour level are circled. These plots are from the same spectrum shown in Figure 4.2

chain ^{1}H and ^{13}C shifts for histidine residues. Aromatic assignments were associated with side-chain data by identifying intraresidue $^{1}H_{\beta}$ to $^{1}H_{\delta}$ NOEs in the ^{13}C-edited NOESY-HSQC spectra. In total, 2038 of 2171 possible ^{13}C, ^{15}N and ^{1}H assignments were obtained for 181 of 182 residues in HRV14-3C protease (94% complete).

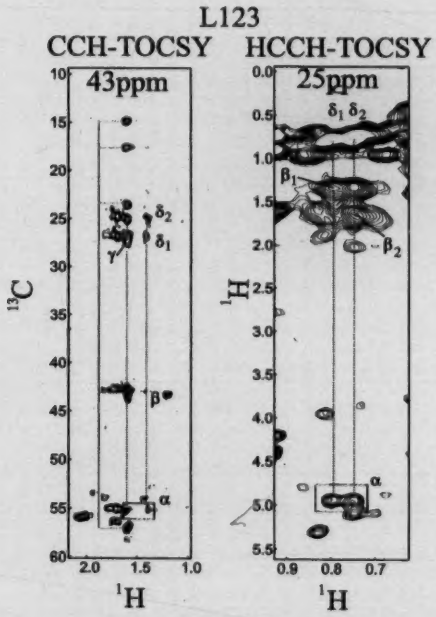

Figure 4.5: **CCH-TOCSY and HCCH-TOCSY Strip Plots for Leu^{123}.** The CCH-TOCSY strip, centered around $^{13}C_{\beta}$ for Leu^{123}, correlates the two geminal $^{1}H_{\beta}$ protons with intra-residue ^{13}C chemical shifts (red lines). The two $^{13}C_{\delta}$ atom chemical shifts (~25 ppm) are resolved. The blue lines connects the overlapping Ile^{37} spin system. The HCCH-TOCSY strip, centered between the $^{13}C_{\delta}$ planes of Leu^{123}, shows the intra-residue ^{1}H-^{1}H correlations connected with red lines.

All possible side-chain proton assignments for the bound ethyl propenoate inhibitor were obtained from the 2D $^{13}C/^{15}N$ F1/F2-filtered TOCSY experiment collected in D_2O. The complete set of ^{1}H, ^{13}C and ^{15}N chemical shift assignments for the inhibitor and inhibitor-bound form of the protein were deposited into the BioMagResBank under accession # 6823. Backbone amide assignments were unattainable for Gly^{1}, Ser^{21}, Asn^{80}, Asn^{110} and Gly^{154}, which occupy the N-terminal, loop and turn regions and present larger than average RMSD values. The complete set of chemical shift assignments are provided in Table D.4.

4.3.2 Inhibited HRV14-3C NMR Structure Calculation

1911 non-redundant, intra-protein NOE assignments (749 short range, 690 medium range, 472 long range) were derived from the ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC [15] experiments. 11 intra-inhibitor and 76 inhibitor-enzyme NOEs were obtained from the F1 filtered, F3 edited-NOESY [16] and the 2D $^{13}C/^{15}N$ F1/F2-filtered NOESY [17] experiments. Structure calculations were improved with the addition of a number of cross β-strand $^{1}H_{\alpha}$-$^{1}H_{\alpha}$ assignments. These assignments were obtained from

a second ^{13}C-edited NOESY-HSQC spectrum collected on the sample dialyzed into D_2O buffer (Figure 4.6).

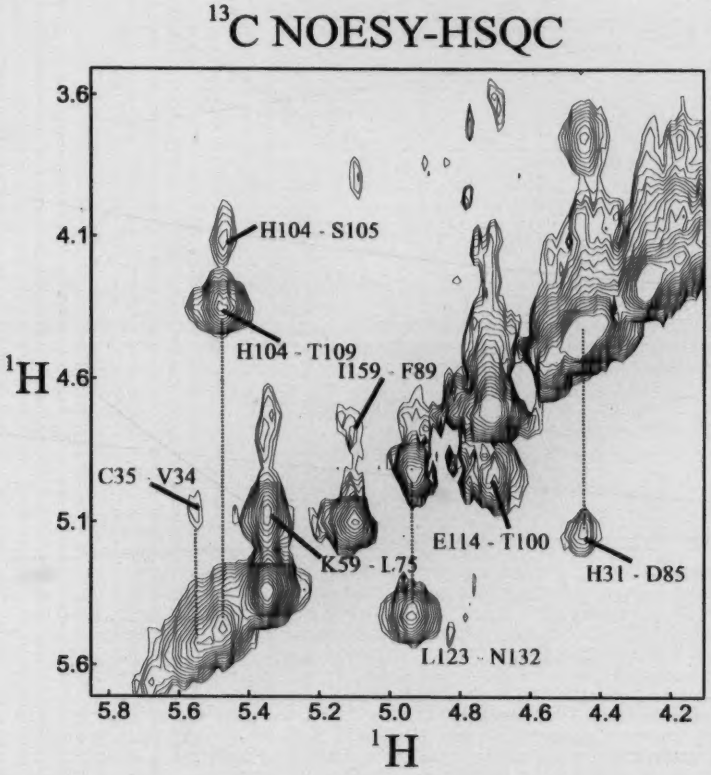

Figure 4.6: **A portion of the ^{13}C-edited NOESY-HSQC Spectrum for the Inhibited HRV14-3C Enzyme**. Water suppression NMR pulses and notch-filter solvent suppression would normally eliminate a large portion of the spectrum around 4.78 ppm. However, a number of cross β-strand $^1H_\alpha$-$^1H_\alpha$ NOEs are visible in this region, including Glu^{114} - Thr^{100}. The [U-^{13}C/^{15}N]-labelled HRV14-3C sample in 99.6% D_2O allowed this spectrum to be collected without water suppression.

The WET water-suppression pulse was not used in the pulse sequence due to the obvious lack of a water signal, and notch-filter solvent suppression was not required in spectral processing. These modifications afforded a number of otherwise unavailable NOE assignments. The 76 protein-inhibitor NOE restraints (obtained from the ^{13}C/^{15}N F1-filtered, F3-edited NOESY experiment) were further cross correlated in the aromatic and aliphatic ^{13}C-edited NOESY experiments. Relatively few NMR signals for the ethyl group of the ethyl propionate ester were identified in either the ^{13}C/^{15}N F1-filtered, F3-edited NOESY, ^{13}C-edited NOESY or

$^{13}C/^{15}N$ -filtered NOESY experiments, which resulted in a relatively large RMSD for the ester group.

NOEs were calibrated using proton cross-peak intensities and binned into three categories with upper bounds of 3.0, 4.0 and 5.5 Å corresponding to strong, medium and weak intensities respectively. All lower bounds were set to 1.8 Å. A total of 131 $^3J_{HNH\alpha}$ coupling constants were unambiguously determined from the HNHA spectrum and used to assign 131 phi (ϕ) angles. In addition, 126 Psi (ψ) angles were predicted using TALOS [18] and SHIFTOR (`http://redpoll.pharmacy.ualberta.ca/shiftor`). The χ_1 for His^{40} was assigned based on H_β to H_N and H_α NOE intensities. The ϕ, ψ and χ angle restraints were assigned with limits of $\pm$ 40°. Hydrogen bonds were identified following analysis of the ^{15}N-HSQC spectra collected on the [U-$^{13}C/^{15}N$]-HRV14-3C protease sample exchanged into 99.6% D_2O buffer. In all, 87 hydrogen bond donors were assigned to amides showing signals 150 minutes following $^1H/^2H$ exchange at 25 °C (pD 6.9) and were given limits of 1.5 - 2.5 Å for H_N to O distances and 2.5 - 3.5 Å for N to O distances. Hydrogen bond acceptors were identified following the initial structure calculations using NOE data alone. Preliminary structures were calculated using the simulated annealing protocol implemented in XPLOR-NIH v2.10 [19, 20]. Inital structure calculations were performed in dihedral space using the PARALLHDG non-bonded parameter set. Center-weighted pseudo-atom corrections were used for ambiguous methylene and methyl proton NOE distances. The preliminary structures generated with these parameters were then used to assign the hydrogen bonds and dihedral angles, which were subsequently used in all remaining structure calculations. A set of 50 structures with minimal violations were chosen from 100 structures initially calculated using the simulated annealing protocol (high temperature steps = 24,000 at 1000 °K; cooling steps = 12,000 with a final temperature of 100 °K). Refinement of these structures was performed with CNS version 1.1 [19] using the RECOORD water refinement protocol [21]. The PARALLHDG and OPLSX non-bonded parameter sets were used for the simulated annealing and water refinement protocols respectively.

An average of 10.5 NOEs per residue were obtained for the structure calculations. 2.6 long-range NOEs per residue and 3.8 medium-range NOE restraints per residue yielded structures with good structural statistics and precision (Table 4.2). Twenty low energy structures presenting good geometry, no improper or dihedral angle violations >5°, no

bond-length violations >0.05 Å and no NOE violations >0.3 Å were selected and deposited in the PDB under accession code 2B0F. The structural statistics of the deposited structures are shown in Table 4.2. AQUA [22] and PROCHECK-NMR [23] were used to calculate and analyze NOE and dihedral angle restraint violations. Six of the 20 deposited structures had a total of 5 NOE restraint violations greater than 0.2 Å. No restraints had violations exceeding 0.3 Å.

Table 4.2: **Structural Statistics for the Inhibited HRV14-3C Protease**

Distance Restraints	
All NOE distances	1998
Intra-residue (protein)	1911
Sequential (| i-j | = 1)	749
Medium (1 <| i-j |≤ 4)	690
Long (| i-j |> 4)	472
Inter-residue (protein-inhibitor)	76
Intra-residue (inhibitor)	11
Hydrogen bonds	87
Violations	
Structures with violations > 0.3Å	0
Structures with violations > 0.2Å	6
Dihedral angle restraints	
All	258
ϕ	131
ψ	126
χ_1	1
Ramachandran Plot†	
Residues in most favored region	78.90%
Residues in additionally allowed region	19.90%
Residues in generously allowed region	0.50%
Residues in disallowed region	0.60%
WHAT-CHECK scores‡	
Second generation packing	-1.25
χ_1 / χ_2	-1.74
RMSD to mean structure§	
Backbone	0.82 ± 0.13
Heavy atom	1.49 ± 0.20
Region 15-78	
Backbone	0.72 ± 0.14
Heavy atom	1.41 ± 0.20
Region 99-103, 111-172	
Backbone	0.56 ± 0.12
Heavy atom	1.17 ± 0.17

†Calculated with PROCHECK-NMR [23]
‡Calculated with WHAT-CHECK [24]
§Calculated with MOLMOL [25]

Geometric and structure quality analysis for the structures were carried out using VADAR [26]. and What-Check [24]. Good geometry was present in the final structures as indicated by What-Check packing and rotamer Z scores following water refinement with RECOORD

[21]. PROCHECK-NMR [23] analysis suggests the solution structure has an equivalent X-ray resolution of 2.4 Å based on Ramachandran plot quality assessment and a resolution of 1.1 Å based on χ_1 pooled and χ_2 trans angle standard deviation assessments. The deposited structures had nearly 80% of their residues in the phi/psi core region. Interestingly, Asp^{32} occupies a disallowed region of the Ramachandran plot (ϕ: 51 $\pm$ 5°, ψ: -91 $\pm$ 5°). Other members of the picornaviridae family, Polio (Asp^{32}), HRV2 (Asp^{32}) and HAV (Asp^{36}) also exhibit similar ϕ and ψ angles (48 $\pm$ 6° and -120 $\pm$ 7° respectively).

The structure calculation of the inhibited HRV14-3C protease involved the incorporation of a previously uncharacterized inhibitor. This required the development of custom XPLOR *topology* and *parameter* files that described the inhibitor. Modifications to XPLOR's peptide *link* files were also required in order to link the inhibitor residues and to covalently attach it to the HRV14-3C enzyme. These files define an unprecedented three-residue *topology* that was needed to link the inhibitor's P_1 glutamine residue and ethyl propionate moiety to the protease's Cys^{146} residue at the S_γ. These files are provided in *Appendix C* and have been included with the PDB deposition.

4.4 Discussion

4.4.1 The Inhibited HRV14-3C Protease Structure

Analysis of medium and long-range NOE patterns [27], $^3J_{HNH\alpha}$ couplings [28, 29] and chemical shifts [30] confirmed α-helical secondary structure for residues Pro^2 - Lys^{12} (**A**) and Thr^{39} - Ala^{41} (**B**) and β-strand secondary structure for residues Ile^{15} - Thr^{20} (**Ia**), Glu^{24} - His^{31} (**Ib**), Val^{34} - Ile^{37} (**Ic**), Asp^{46} - Leu^{48} (**Id**), Gln^{52} - Asp^{64} (**Ie**), Leu^{72} - Arg^{79} (**If**), Ala^{99} - Val^{103} (**IIa**), Val^{115} - Leu^{126} (**IIb**), Pro^{130} - Arg^{136} (**IIc**), Val^{149} - Ala^{152} (**IId**), Lys^{155} - Gly^{163} (**IIe**), and Arg^{166} - Gln^{172} (**IIf**). These elements are folded into two six-stranded β-barrel domains (Figure 4.7) that accommodate the active site residues, His^{40}, Glu^{71} and Cys^{146}, between them in a shallow cleft. The RNA binding site, $KFRDI^{82-86}$, is capped by a short 3_{10} helix (Arg^{87} - Phe^{89}) and resides in a random coil that tethers these domains. This coil is on an opposite side of the protein relative to the proteolytic site. These structural elements are conserved among all the picornaviridae 3C protease structures reported to date.

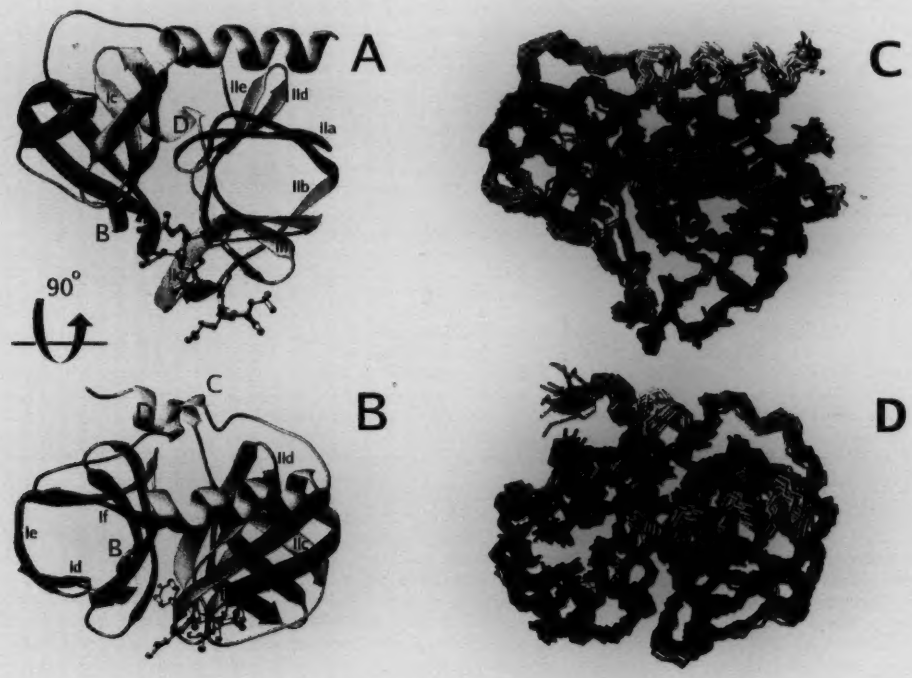

Figure 4.7: **Ribbon Representation of the Inhibited HRV14-3C Protease**. α-helices are colored yellow and labelled with capital letters. β-strands are colored orange and labelled with lower case subscript. The N-terminus β-barrel domain has the prefix I and the C-terminus β-barrel has the prefix II. Representation **B** is rotated 90° about the x-axis with respect to representation **A**. The acetyl-LEALFQ-ethyl propionate inhibitor is rendered as a ball and stick. **C**. Structural ensemble of the HRV14-3C protease with the C-terminal β-barrel domain superposed (residues 99 - 103 and 111 - 172: backbone RMSD ~0.56 Å). **D**. Structural ensemble of the HRV14-3C protease with the N-terminal β-barrel domain superposed (residues 15 - 78: backbone RMSD ~0.72 Å). Aromatic side-chains are colored red and branch chain amino acid side-chains are colored blue in the structural ensembles.

The solution structure of inhibited-HRV14-3C has been solved with an overall backbone atom RMSD of 0.82 ± 0.13 Å and 1.49 ± 0.20 Å for all heavy atoms from residues Gly^{1} to Glu^{180}. A larger RMSD over the termini and loop regions (residues Asp^{64} - Ile^{68} and His^{104} - Asn^{110}) skewed the global RMSD to larger values compared to the β-barrel domains alone.

4.4.2 Induced Fit Substrate Recognition

The C-terminal β-barrel (Ala^{99} - Leu^{173}) is involved with the inhibitor binding and provided a slightly tighter RMSD (0.56 ± 0.12 Å for backbone atoms and 1.17 ± 0.17 Å for all heavy atoms) compared with the β-barrel formed between residues Ile^{15} to Asp^{78} (0.72 ± 0.14 Å for backbone atoms and 1.41 ± 0.20 Å for all heavy atoms).

In order for the 3C enzyme to recognize and bind the various natural substrates, it is suspected that localized conformation changes within the substrate binding interface must occur. To investigate whether changes occur in any other part of the protease following inhibition, the protein's solvent accessible surface area changes were compared with chemical shift perturbations upon inhibition. The surface area changes were calculated with the program STC [31] and the chemical shift changes were calculated with the program CSDIFF (`http://www.bionmr.com/csdiff`). The perturbation analysis used all available backbone chemical shift data available for the apo (BMRB # 5659) and the inhibited (BMRB # 6823) HRV14-3C proteases (Figure 4.9) as this data is sensitive to backbone torsion angles and could identify both local and global changes. These results show that backbone chemical shift changes do not occur in any region of the protease which is not involved with inhibitor binding. There are specific data, however, that suggest localized

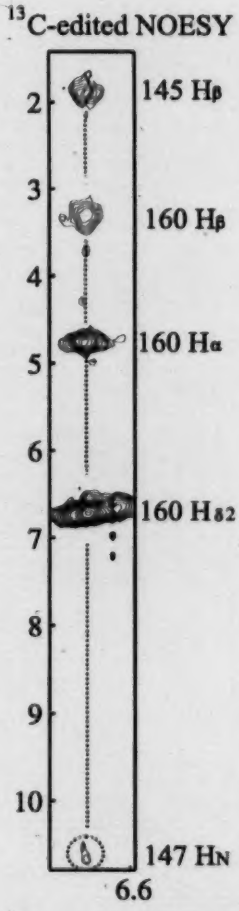

Figure 4.8: **^{13}C -NOESY-HSQC Spectra for His^{160}'s $^{1}H_{\delta 2}$.** NOE data for the $^{1}H_{\delta 2}$ atom of His^{160} shows the long range NOE to the H_N atom of Gly^{147}.

conformational changes occurs. First, a 2 ppm downfield shift is observed for the ^1H_N proton of Gly^{147} upon inhibition, which indicates a reduction in hydrogen bond length [32] to His^{160}'s carbonyl oxygen. Gly^{147} is sequentially linked to the active site cysteine (Cys^{146}), however, makes no direct contact with the acetyl-LEALFQ-ethyl propionate inhibitor. Furthermore Gly^{147}'s amide proton is oriented into the C-terminal β-barrel domain and forms anti-parallel β-strand hydrogen bonds with His^{160}. This latter residue is important for P_1 substrate recognition and binding. The orientation of Gly^{147}'s backbone amide in the inhibited form was confirmed with NOE data for the $^1H_{\delta 2}$ proton of His^{160} (Figure 4.8). This was critical because overlapped peaks presented in the ^{15}N-NOESY-HSQC for Gly^{147}'s and His^{160}'s ^1H_N chemical shifts (10.64 ppm and 10.47 ppm respectively).

In addition to Gly^{147}, chemical shift changes are observable for other backbone atoms not making direct contact with the inhibitor. Examples include Ser^{170} and Ala^{171} (Figure 4.9). These residues are localized between the two β-barrel domains and form β-strand hydrogen bonds with Ile^{135} and Ile^{159} respectively. Interestingly, backbone amide chemical shift assignments could not be made for Gln^{172} in the apo enzyme, which forms β-strand hydrogen bonds with Arg^{133} in the bound state. Presumably, this results from intermediate chemical exchange in the apo form. All of these residues are localized on the interior face of the C-terminal β-barrel domain and are positioned like rungs on ladder (Figure 4.10). These data indicate that conformational changes, whether local or global domain repositioning, occur upon substrate binding. Although these findings reflect observations made for other picornaviral 3C enzymes [33], they are the first findings for any member of the rhinovirus family.

Additionally, larger than average RMSDs are observed for the $\mathbf{II}_b$, $\mathbf{II}_c$ and $\mathbf{II}_e$ β-strands in the homologous apo Polio-3C protease (Figure 3.1). These β-strands are involved in substrate interactions. In the apo form of the HRV14-3C protease (BMRB # 5659), these regions along with the residues involved in inter-domain contacts (Leu^{70}, Leu^{72}, Arg^{133}, Met^{134}, His^{160}, Ala^{171} and Gln^{172}) present exchange broadened amide signals. These findings further support the hypothesis that conformational flexibility and localized induced fitting of the enzyme may be required to recognize the various natural peptide substrates.

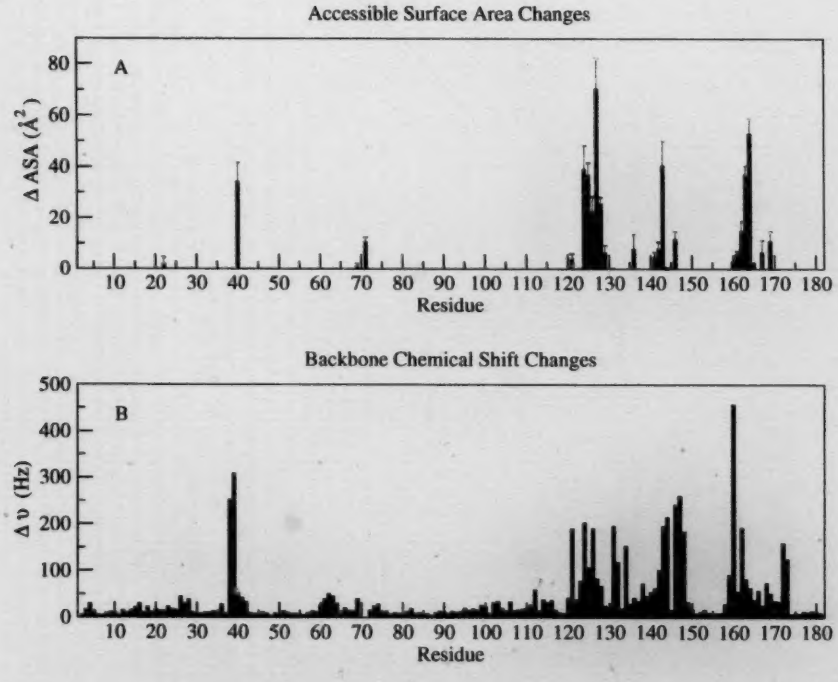

Figure 4.9: **Chemical Shift and Solvent Accessible Surface Area Changes Upon Inhibition of the HRV14-3C Protease**. **A**. Surface areas changes for the deposited structures (PDB code 2B0F) were calculated with STC [31]. **B**. The chemical shift changes at 500 MHz were calculated with CSDIFF (`http://www.bionmr.com/csdiff`) using the formula: $\frac{|\nu_{HN1}-\nu_{HN2}|+|\nu_{N1}-\nu_{N2}|+|\nu_{C\alpha 1}-\nu_{C\alpha 2}|+|\nu_{C\beta 1}-\nu_{C\beta 2}|+|\nu_{C'}-\nu_{C'}|}{\#\ atoms\ used}$. A correlation with the solvent accessible surface area changes upon binding of the inhibitor exists. These chemical shift changes are localized to residues within the C-terminal β-barrel domain in close proximity to the inhibitor.

4.4.3 Active Site Triad Comparison

There are two X-ray crystallographic structures of 3C proteases that exhibit good sequence similarity to the HRV14-3C protease. One is the X-ray structure of the HRV2-3C protease bound to the AG7088 inhibitor (PDB code 1CQQ) reported by Matthews *et al.* [34] and the other is the apo Polio-3C protease (PDB code 1L1N) reported by Mosimann, *et al.* [35]. These proteins share 51% and 47% sequence identity, respectively, with the HRV14-3C enzyme. Comparison between the average inhibited HRV14-3C protease structure and these homologous enzymes yield backbone RMSDs of 1.32 Å and 2.76 Å respectively. Analysis

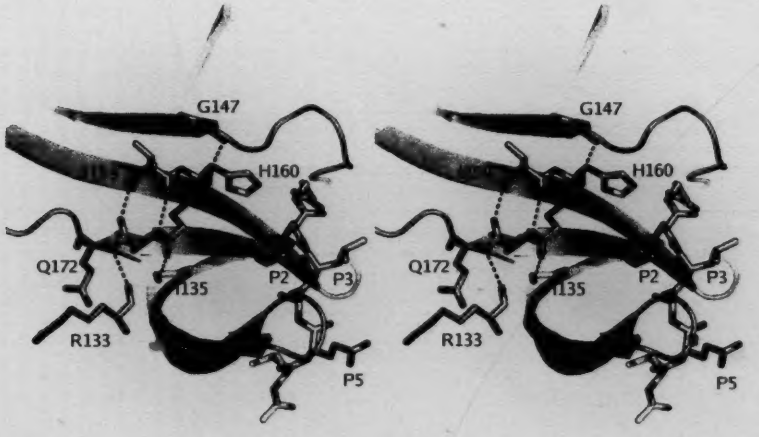

Figure 4.10: **Hydrogen Bonds Within the C-terminal β-barrel Domain.** A wall-eyed stereo view of the hydrogen bond network among stacked β-strands in the C-terminal β-barrel domain of the inhibited HRV14-3C protease. These residues are sandwiched between the two β-barrel domains and do not make direct contact with the surface-interacting inhibitor. The position of the acetyl-LEALFQ-ethyl propionate inhibitor relative to these residues is also depicted. Because the inhibitor it is far enough away from these residues (> 5 Å), it cannot have any direct influence on their chemical shifts.

of the natural and peptide-mimic ethyl propionate inhibitors bound to the HRV14-3C and HRV2-3C proteases respectively, reveal that the two inhibitors bind in a similar fashion up to the P_3 position. In all these structures, the amides of the enzymatic residues preceding the catalytic site (Gly^{144}, Gln^{145} and Cys^{146} in the HRV14-3C protease) produce a net positive charge that creates the oxyanion hole (Figure 4.11). This '*hole*' presumably stabilizes the cleaved carbonyl following proteolysis. In the inhibited HRV14-3C protease, this hole is occupied by the H_1 proton of the ethyl propionate moiety. This proton is bound to the carbon forming a covalent bond with the Sγ of Cys^{146}. In the natural peptide substrate this position would correspond to the carbonyl of the P_1 residue, which would be stabilized by the oxyanion hole upon peptide cleavage. Hydrogen bonding between the corresponding Cys^{147} and Gly^{145} amides within this oxyanion hole and the carbonyl oxygen of the ethyl propionate ester was observed in the X-ray crystallographic study of the inhibited HRV2-3C protease [34]. In the HRV2-3C protease study, this oxygen resides equidistant (~2.23 Å) between the two amide hydrogens.

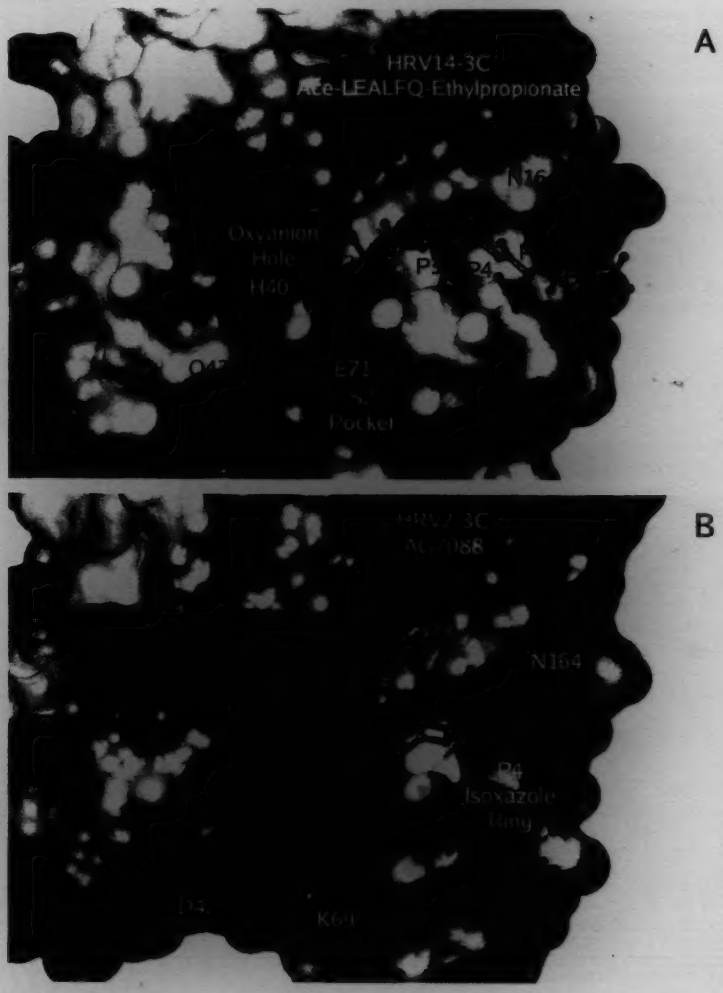

Figure 4.11: **Electro Molecular Surfaces of the HRV14-3C and HRV2-3C Proteases.** (**A**): The acetyl-LEALFQ-ethyl propionate inhibitor is rendered as a ball and stick figure on the electrostatic molecular surface of the HRV14-3C protease. (**B**): The peptide-mimic inhibitor, AG7088, is rendered as a ball and stick figure on the electrostatic map of the HRV2-3C protease. In both figures, negative charges are colored red and positive charges are colored blue. The oxyanion hole that stabilizes the P_1 carbonyl and binds the post-scissile peptide is positively charged. The back of the deep S_2 pocket accommodates a variety of amino acids and is negatively charged from the side-chain carboxylate group of the active site Glu^{71}. The P_2 phenylalanine of the acetyl-LEALFQ-ethyl propionate inhibitor that occupies this pocket makes van der Waals contacts with Leu^{126} and Thr^{131} of the II_b and II_c β-strands respectively. Electrostatic differences of the S_2 pocket between the different enzymes result from D42Q and K69N sub-genus substitutions.

However, in our study, we found the ester of the inhibitor occupying two conformations that directed the corresponding oxygen (OX_8) toward the amides of Cys^{146} and Gly^{144} with distances of 2.36 ± 0.35 Å and 2.29 ± 0.18 Å or distances of 3.26 ± 0.21 Å and 1.99 ± 0.26 Å respectively. These orientations result from the lack of NOE and hydrogen bond restraints for the ethyl propionate group. Although, the lack of data could result from intermediate conformation exchange, it should be noted that unlike Gly^{163}, neither Cys^{146} nor Gly^{145} presented amide signals in our deuterium exchange experiments despite all of these amides being solvent exposed and involved with substrate interactions.

There are subtle differences between the active site catalytic triad residues (cysteine, histidine and glutamic acid) of the solution HRV14-3C and the X-ray HRV2-3C proteases. Furthermore, the orientation of the catalytic triad residues in the inhibited enzymes resembles that of the residues in the apo Polio-3C protease.

Comparison of the active site residues among the different 3C proteases indicates that the orientation of the His^{40} and Glu^{71} side-chains in HRV14-3C protease are similar to the reported structures of the apo form described by Matthews *et al.* [36] and the deposited structures for the HRV2-3C [34] and the Polio-3C [35] proteases. In all cases the side-chain carboxylate group from the active site glutamic acid (Glu^{71}) forms a salt bridge with the side-chain imidazole $N_{\delta 1}$ atom from the active site histidine (His^{40}), which allows the imidazole's $N_{\epsilon 2}$ to direct toward the $S\gamma$ of Cys^{146} thereby forming the proteolytic acid/base catalyst. This orientation is shown in Figure 4.12.

However, a discrepancy regarding the protonation state of the His^{40} side-chain exists between the X-ray and NMR structures. Analysis of the histidine side-chain $C_{\delta 2}$ and $C_{\epsilon 1}$ chemical shifts in the HRV14-3C protease revealed values of 117.39 ± 1.5 ppm for the $C_{\delta 2}$ shift and 138.39 ± 2.32 ppm $C_{\epsilon 1}$ shift at pH 6.5. These data suggest a single protonation state [37], which disagrees with the double protonated state suggested for the His^{40} residue deposited with the HRV2-3C protease structure (PDB code 1CQQ). Furthermore, both structures were solved at pH 6.5, which is near the pK_a where neutral tautomeric states of histidine's side-chain exist.

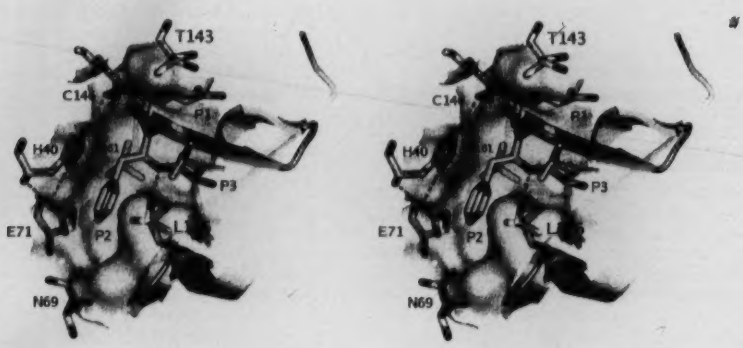

Figure 4.12: **Catalytic Triad and P_1 to P_3 Substrate Interactions**. Wall-eyed stereo view of the ethyl propionate group and P_1 - P_3 substrate residues. A large cleft between the β-barrels is occupied by the P_2 phenylalanine ring, which makes hydrophobic contacts with Leu^{126} and Val^{161} (colored yellow). The P_3 leucine is large enough to fill the substrate cleft and make van der Waals contacts with the γ methyl group on Thr^{143}. Also shown in the side-chain of Ser^{127} that acts as a hydrogen bond acceptor to the P_2 amide in ~50% of the calculated structures.

4.4.4 Substrate Binding: Backbone and Sidechain Contacts

The substrate used in this study involved six "***naturally***" occurring residues (LEALFQ) in contrast to the four "unnatural" peptide-mimic residues of AG7088. The added length of our inhibitor allowed for a detailed characterization of interactions between the enzyme and substrate P_3 to P_6 residues.

Hydrogen Bond Interactions

Our results indicate that the peptidyl inhibitor binds anti-parallel to the $\mathbf{II}_b$ and $\mathbf{II}_e$ β-strands of HRV14-3C. The amide of the P_1 glutamine acts as a hydrogen donor to the carbonyl of Val^{161} and hydrogen bonds are formed between the backbone atoms of the P_3 and Gly^{163} residues. Our study also indicates that the backbone atoms of the P_4 residue bridge the $\mathbf{II}_b$ β-strand and form hydrogen bonds with the backbone amide of Ser^{127} and the backbone carbonyl of Asn^{125}. The side-chain of Asn^{164} orients to act as either a hydrogen bond donor or acceptor to the corresponding backbone atoms of the P_5 residue. Finally, the backbone carbonyl of the P_6 residue acts as a hydrogen bond acceptor for the backbone

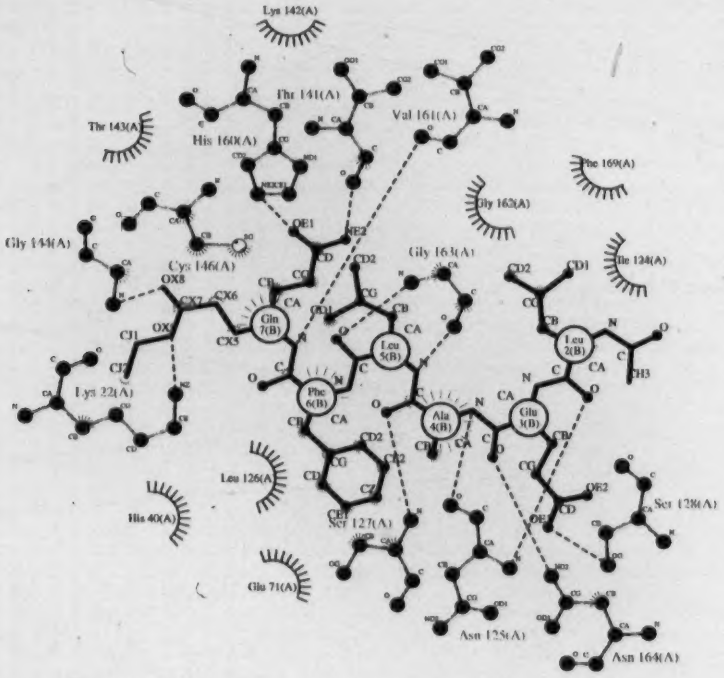

Figure 4.13: **Schematic Representation of the Pharmacophore Interactions**. The representation was created with LIGPLOT [38]. The network of hydrogen bonds between the peptide substrate (segment B) and the enzyme depicts an anti-parallel β-strand formation (represented as broken lines). Other key hydrogen bonds and van der Waals contacts with conserved HRV-3C residues are also shown. The covalent bond between Cys^{146} and the CX_5 carbon of the inhibitor is not shown.

Asn^{125} amide. These interactions are shown in Figure 4.13. $^1H/^2H$ exchange data confirmed the hydrogen bond donors of Gly^{163} and Ser^{127}. No deuterium exchange data presented for the side-chain amide of Asn^{164} whose orientations are a reflection of structure calculation refinement.

S_1 to S_3 Substrate Pockets

In addition to these hydrogen bond interactions, the side-chain of the P_1 glutamine makes a number of van der Waals contacts by filling a shallow cleft between the the coil region formed between residues Thr^{141} to Gly^{144} and the $\mathbf{II}_e\beta$-strand residues His^{160} to Gln^{164}. The preferred P_1 glutamine positions its side-chain amide to form three key hy-

drogen bonds: one from its carbonyl to the $H_{\epsilon 2}$ of His^{160} and the other two from its amide to the backbone carbonyl and side-chain hydroxyl of Thr^{141}. The ϕ angle of the conserved threonine residue (HRV14-3C: Thr^{141}; HRV2-3C Thr^{142}) differs between the two enzymes (HRV14-3C: -114 $\pm$ 15°; HRV2-3C: -57°). This probably originates because the threonine in the HRV2-3C protease is orientated outward in order to form a hydrogen bond with the imine group of the 2-oxo-pyrrolidin-3-yl ring of AG7088. The aromatic ring of the P_2 phenylalanine and the branched side-chain of P_3 (a leucine) make van der Waals contacts with Leu^{126} and Ser^{127} (Figure 4.12). The P_3 leucine is solvent exposed and covers the substrate binding pocket spanning the P_1 and P_2 residues.

S_4 to S_6 Substrate Pockets

The methyl group of the alanine in the P_4 position is tucked inside the S_4 binding pocket and makes contact with the $^1H_\alpha$ protons of Gly^{163} and the aromatic protons of Phe^{169}. The P_5 glutamic acid side-chain is oriented parallel to the II_b β-strand backbone and makes van der Waals contacts with the side-chain of Asn^{125} and forms a hydrogen bond with the hydroxyl of Ser^{128}. The side-chain of the P_6 leucine is directed away from P_5 side-chain and fills the remainder of the binding pocket below Asn^{164} making numerous van der Waals contacts with the aromatic ring of Phe^{169} and the side-chain of Ile^{124}.

4.4.5 Substrate Specificity and Pharmacophore Analysis

P_1 Substrate Binding

The specificity of the HRV14-3C protease for a glutamine residue in the P_1 position has been previously demonstrated. the three hydrogen bonds that were previously observed between AG7088 and the HRV2-3C protease are mimicked again in the peptidyl inhibitor bound to the HRV14-3C protease. Despite satisfying two of the three hydrogen bonds with a Q/E substitution (1C/1D cleavage sequence in the HRV polyprotein), a dramatic cleavage rate reduction is observed [39, 40]. This change in affinity might be attributed to mutually repulsive forces between the carboxylate of the substituted glutamic acid and the backbone carbonyl of Thr^{141} due to their proximity upon binding.

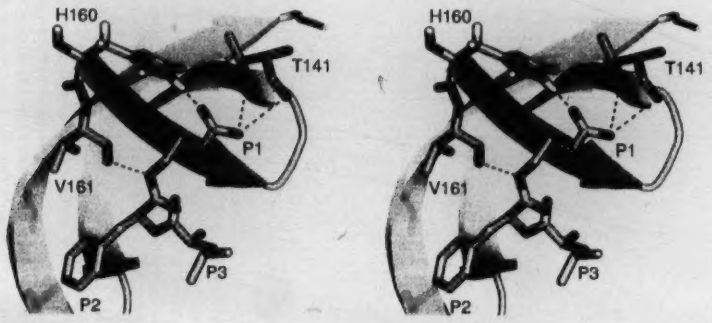

Figure 4.14: **P_1 and P_2 Substrate Interactions**. Wall-eyed stereo view of the numerous P_1 and P_2 hydrogen bond interactions, including the three bonds formed with conserved 3C protease amino acids His^{160} and Thr^{141}. Some protons have been removed for clarity.

P_2 Substrate Binding

The S_2 pocket is a large, deep, negatively charged cleft between the two β-barrel domains (Figure 4.11) that presents variable substrate recognition. Despite the negative electrostatic charge that results from the presence of the active site Glu^{71} in the back of the pocket, a number of synthetic inhibitors have incorporated aromatic and poly-aromatic rings at the P_2 position with favorable results [41, 42]. The P_2 phenylalanine used in our inhibitor bound in a similar fashion to the fluoro-substituted aromatic ring incorporated in AG7088, whereby its aromatic ring significantly reduces the protein's solvent accessible surface area and makes numerous van der Waals contacts with the enzyme (His^{40}, Asn^{69}, Glu^{71}, Leu^{126}, Ser^{127}, Thr^{129}, Thr^{131} and Val^{161}). These interactions are shown in Figure 4.12. The variable recognition between the rhinovirus serotypes possibly stems from the neutral to charged residue substitution observed in this pocket on the N-terminal β-barrel domain. The neutral residues, Gln^{42} and Asn^{69}, found in subgenus B rhinoviruses differ from the charged residues, Asp^{42} and Lys^{69}, found in subgenus A rhinoviruses. These changes affect the electrostatic surface of the binding pocket (Figure 4.11) and may account for the increased substrate recognition that the HRV14-3C protease demonstrates when hydrophobic groups are incorporated in the P_2 position [42]. The backbone amide of the P_2 residue acts as a hydrogen bond donor for the hydroxyl of the highly conserved Ser^{127} residue [43]. However, only 55% of the calculated solution structures exhibited this hydrogen bond while the

remaining structures had the hydroxyl group orientated into the solvent. This structural ambiguity was observed in previously published structure-activity work whereby omission of the P_2 backbone amide bond was adopted, which did not negatively impact substrate recognition [42]. However, it should be noted that the side-chain orientations of Ser^{127} observed in the NMR ensemble result from the structure calculations as hydroxyl protons exchange in solution and are unobservable in NMR, therefore, no restraint information was obtainable.

P_3 Substrate Binding

The HRV14-3C protease demonstrates a preference for a larger branch-chain amino acid (leucine) in the P_3 position, in contrast to the smaller branch-chain amino acid (valine) preferred for subgenus A rhinovirus serotypes [42]. This larger residue was incorporated in our study to analyze this difference. Sub-genus diversity appear to stem from van der Waals contacts afforded between the δ methyl group of the P_3 leucine and the γ methyl group of Thr^{143} (Figure 4.15). A threonine residue is found in the sub-genus B rhinovirus 3C proteases, whereas this residue is substituted for the smaller serine in sub-genus A rhinovirus 3C proteases [43]. The leucine residue is also of sufficient size to span the entire S_3 pocket and make contact with the P_1 glutamine and P_2 phenylalanine residues.

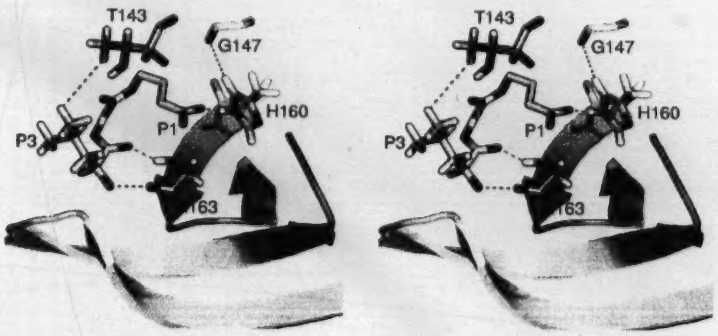

Figure 4.15: **P_3 Substrate Interactions**. Wall-eyed stereo view of the P_3 leucine within the S_3 binding pocket. Hydrogen bonds are shown with dashed orange lines. Key NOEs, including the long range NOE between Gly^{147} H_N and His^{160} $H\delta_1$, are shown with dashed, dark blue lines. The P_3 $^1H_\delta$ and HRV14-3C Thr^{143} $^1H_\gamma$-methyl groups are within 5 Å and make van der Waals contacts. The backbone hydrogen bonds between the P_3 leucine and Gly^{163} are clearly visible.

P_4 to P_6 Substrate Binding

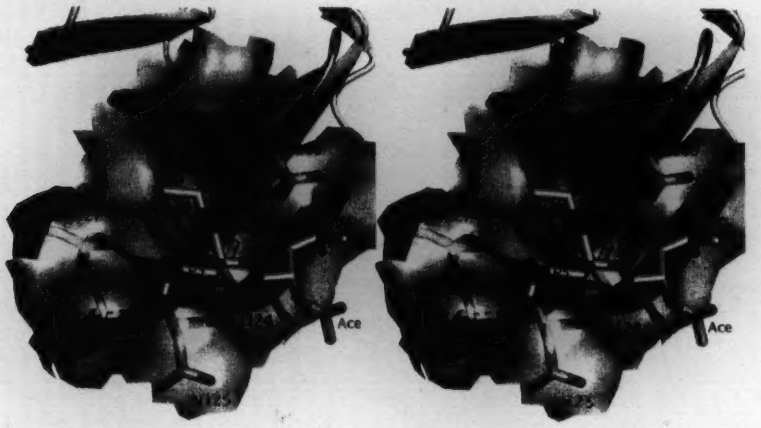

Figure 4.16: **P_4 to P_6 Substrate Interactions**. Wall-eyed Stereo view of the S_4 - S_6 substrate pocket. A number of hydrogen bonds between the inhibitor and enzyme are shown as dashed, black lines. The hydrophobic side-chains of Phe^{169} and Ile^{124}, colored yellow, can be seen making van der Waals contacts with the methyl group of the P_4 alanine and side-chain of the P_6 leucine residue.

Substrate cleavage studies have demonstrated that inclusion of the P_5 residue results in a 4-fold increase in substrate recognition, while inclusion of the P_6 to P_8 residues increases substrate cleavage a further 2-fold (relative K_{cat}/K_m) [44]. Analysis of this phenomenon with our substrate was conducted with the program STC [31]. This analysis revealed that elimination of the P_6 leucine residue increased the enzyme's non-polar solvent accessible surface area by 45 ± 15 Å^2 and predicted a reduction of the K_D by 24 ± 5-fold. These results support the experimental observations that shortened substrates have impaired recognition [45].

No residue specificity for the P_5 amino acid has yet been identified. As such, this residue is solvent exposed and makes contact with hydrophilic residues (Asn^{125} and Ser^{128}). The hydrophobic leucine residue that occupies the P_6 position of the inhibitor folds into the shallow S_6 binding cleft under Asn^{164} and makes a number of van der Waals contacts with the side-chains of Ile^{124} and Phe^{169}. These hydrophobic contacts explain the serotype conservation of these 3C protease residues [43] and the preference for a hydrophobic residue in this position.

The conserved Asn^{164} residue (Table 4.3) is well positioned to interact with the P_4 to P_6 residues. Interestingly, this asparagine residue also undergoes deamidation, which dramatically reduces substrate recognition by 10-fold [46]. It is still not known if the deamidation mechanism is a host-cell mediated, autocatalytic or a non-enzymatic event. It is known that having a glycine following an asparagine residue can facilitate a spontaneous deamidation event of asparagine residues. However, to date this reaction has not been investigated nor has it been characterized in any other 3C protease. It is noteworthy that the majority of other 3C proteases have conserved Asn-Gly sequences, including HRV2-3C, Polio-3C and FMD-3C. Regardless of the exact mechanism, the impact this deamidation event has on substrate recognition and therefore, possible auto-regulatory mechanisms are evident. Therefore, to help identify the specific enzyme-substrate binding interactions in this region of the protease we included the P_5 and P_6 residues within the ethyl propionate peptidyl inhibitor.

HRV14-3C's Asn^{164} residue is incorporated within a β-turn between the $\mathbf{II}_e$ and $\mathbf{II}_f$ β-strands. Sequentially, it is positioned at the '*i*' position of a β-turn for both the HRV and Polio 3C proteases and at the '*i+1*' position for the HAV and FMD 3C proteases (Table 4.3). Our study indicates its side-chain is important for forming hydrogen bonds with the P_5 backbone atoms and contributing to the anti-parallel binding orientation observed between the substrate and the $\mathbf{II}_b$ and $\mathbf{II}_c$ β-strands . Our structure calculations produced two orientations of the chi angles for Asn^{164} (χ_1: 172 $\pm$ 5° and χ_2: -90 $\pm$ 2° or χ_1: 177 $\pm$ 5° and χ_2: 33 $\pm$ 2°), which direct its side-chain amide or carbonyl toward either the backbone carbonyl or amide of the P_5 substrate residue respectively (Figure 4.17).

This was surprising as the side-chain was hypothesized to form anti-parallel type hydrogen bonds with both the backbone amide and carbonyl groups of the P_5 residue, thereby perpetuating the anti-parallel β-strand hydrogen bond network observed with upstream residues. However, this alternate orientation was confirmed from NOEs observed between the $^1H_\beta$ protons of Asn^{164} and the $^1H_\alpha$ and $^1H_\beta$ protons of the P_4 alanine. These orientations of the side-chain amide group, however, result solely from structure refinement as no NOE or hydrogen bond restraints could be assigned to the exchangeable amide protons. The former orientation described allows for an electrostatic interaction between the Asn^{164} side-chain amine and the P_5 backbone carbonyl. The angle formed between the amide's

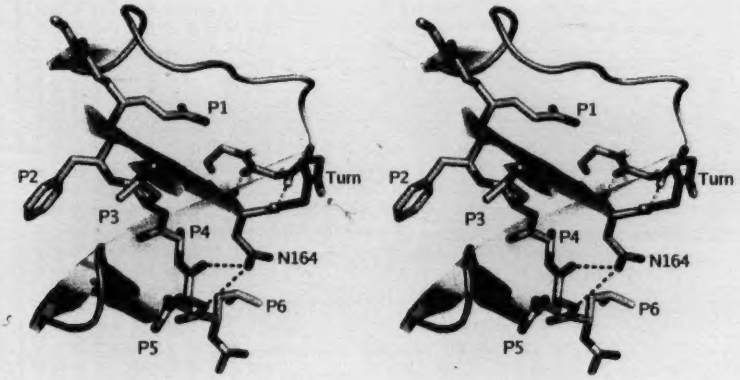

Figure 4.17: **Substrate Interactions with HRV14-3C's Asn^{164} Residue.** Wall-eyed stereo view of the substrate P_5 backbone atoms with the side-chain of Asn^{164}: χ_1: 172° and χ_2: -90° . The two possible hydrogen bonds formed from the two orientations are shown as broken black lines. Yellow broken lines define the hydrogen bonds seen in a typical anti-parallel β-turn.

nitrogen and H_N atoms and the interacting carbonyl oxygen is ~97°, which does not define an ideal hydrogen bond angle between anti-parallel β-strands. Furthermore, this orientation positions Asn^{164}'s side-chain and backbone carbonyl groups within 3 Å of each other, which could result in electrostatic repulsion. The latter orientation, however, allows for the formation of two hydrogen bonds. First, the side-chain ^1H_N of Asn^{164} forms a hydrogen bond with its own backbone carbonyl group, and second, its side-chain carbonyl accepts a hydrogen bond from the backbone ^1H_N of the P_5 residue. The hydrogen bond angle formed from this interaction is ~120°, which is closer to the ideal "linear" orientation observed for atoms involved in anti-parallel β-strand hydrogen bond interactions. Furthermore, this orientation would not suffer from electrostatic repulsion and would form an additional hydrogen bond.

Differences exist between the two known inhibited HRV-3C proteases with respect to their backbone torsion angles and inhibitor interactions in this region as well (Table 4.3). Hydrogen bonding is observed between the conserved Asn^{165} of HRV2-3C and the isoxazole ring of the peptide-mimic inhibitor AG7088, which occupies the P_4 position. However, to accommodate the inhibitors isoxazole ring that sits orthogonal to the β-turn residues, the $\mathbf{II}_e/\mathbf{II}_f$ β-strands are pushed upward and Asn^{165} adopts χ_1 and χ_2 angles of 60° and 30°

respectively in order to allow this hydrogen bond to form (Figure 4.18). This orientation is stabilized through hydrogen bonds formed from Asn^{165}'s side-chain carbonyl and the *i+2* and *i+3* backbone $^{1}H_{N}$ atoms. In contrast, the HRV14-3C protease forms well-defined type II′ β-turn backbone torsion angles while interacting with the natural peptide substrate (Figure 4.17). Well-defined β-turn angles are also observed for homologous 3C enzymes without inhibitors or with small inhibitors that do not interact with these residues (Table 4.3).

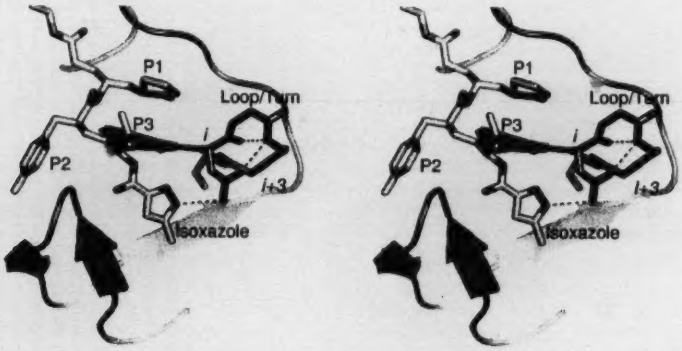

Figure 4.18: **Substrate Interactions with HRV2-3C's Asn^{165} Residue.** Wall-eyed stereo view of Asn^{164}'s side-chain interactions with AG7088's isoxazole ring. Hydrogen bonds are shown as broken orange lines. The anti-parallel β-strands are pushed up to accommodate the isoxazole ring. This orientation is stabilized with hydrogen bonds from the side-chain carbonyl of Asn^{165} to the *i+2* and *i+3* backbone amide atoms.

Table 4.3: **Backbone angles of the Picornaviridae 3C^{pro} II$_e$ - II$_f$ Loop/Turn Region**

PDB		Residues[†]	$i\ \phi$	$i\ \psi$	$i+1\ \phi$	$i+1\ \psi$	$i+2\ \phi$	$i+2\ \psi$	$i+3\ \phi$	$i+3\ \psi$	Turn[‡]
HRV14-3C[§]	2B0F	N^{164}-G,R-Q^{167}	-132 ± 5	96 ± 10	69 ± 4	-107 ± 8	-101 ± 12	8 ± 12	-119 ± 10	125 ± 11	II′
HRV2-3C[§]	1CQQ	N^{165}-G,R-Q^{168}	-130	27	89	-27	-102	-62	-106	161	-
POLIO-3C	1L1N	N^{165}-G,S-Q^{168}	-114 ± 15	60 ± 27	69 ± 10	-35 ± 30	-122 ± 23	-28 ± 5	-118 ± 1	152 ± 11	-
HAV-3C	1HAV	G^{195}-N,S-I^{198}	-148 ± 10	-90 ± 21	-75 ± 22	-73 ± 10	-150 ± 21	24 ± 6	-124 ± 1	134 ± 3	-
HAV-3C[§]	1QA7	G^{195}-N,S-I^{198}	159 ± 28	149 ± 4	60 ± 3	42 ± 8	61 ± 11	16 ± 10	-120 ± 8	137 ± 8	III′
HAV-3C[§]	2A4O	G^{195}-N,S-I^{198}	168	152	51	42	60	30	-123	130	III′
HAV-3C[§]	2CVX	G^{195}-N,S-I^{198}	145	155	53	37	59	30	-130	132	III′
FMD-3C	2BHG	G^{185}-N,G-V^{188}	-179 ± 5	154 ± 6	47 ± 2	41 ± 5	92 ± 3	-4 ± 4	-132 ± 3	155 ± 9	I′

[†]Superscript annotation indicates i to $i+3$ sequential position
[‡]β-turn ϕ and ψ angles. Type I′: $i+1$ 60,30 $i+2$ 90,0; Type II′: $i+1$: 60, -120 $i+2$ -80,0; Type III′: $i+1$ 60,30 $i+2$ 60,30
[§]Structures with bound inhibitors.

Post-Scissile Bond Interactions

Despite evidence that only residues down-stream from the scissile bond are necessary for substrate recognition and cleavage [44], inclusion of Gly and Pro residues in the P'_1 - P'_2 positions has been shown to increase substrate recognition significantly [39, 40]. Furthermore, SAR studies of the HRV-3C proteases with inhibitors indicate that significant improvements in binding affinity is afforded by incorporating aromatic moieties that occupy the P'_1 and P'_2 position [47]. These results parallel those obtained by Jewel and coworkers [48] who investigated the cleavage rates of natural peptide substrates for the HAV-3C protease and found that inclusion of a phenylalanine in the P'_2 position increases substrate recognition by 3 to 4-fold. Although no structural data for the S'_1 - S'_3 binding pockets of the HRV proteases exists, some hypotheses regarding post-scissile interactions can be made based on the data obtained from the solution structure of the inhibited HRV14-3C protease. In particular, the positively charged S'_1 and S'_2 pocket is lined with hydrophilic residues (Lys^{22}, Glu^{24}, Ser^{105}, Asn^{106} and Asn^{107}). This pocket could accommodate the conserved P'_1 and P'_2 residues (glycine and proline) provide favorable electrostatic and hydrogen bond interactions. The P'_3 hydrophobic residue would then be positioned close enough to interact with the surface-exposed Phe^{108} residue that resides in the loop/turn region between $\mathbf{II}_a$ and $\mathbf{II}_b$. This region exhibits conformational flexibility in the solution structure, presumably from the lack of post-scissile substrate interactions. Further evidence that supports this idea a are the exchange broadened NMR signals observed in this region. These data are iterated by the large B-factors observed in the substrate binding region for the homologous apo Polio-3C protease (Figure 3.1). Interestingly, in addition to the aromatic phenylalanine residue, the S'_3 pocket also contains the hydrogen bond forming side-chain from Gln^{145}. These residues are conserved in the sub-genus A rhinoviruses, however, they are spatially flipped (HRV2-3C: Gln^{108} and Tyr^{146}).

4.5 Conclusion

The multiple rhinovirus serotypes present sub-genus diversity regarding substrate recognition. This diversity has presented some barriers in developing a single broad spectrum therapeutic to treat rhinovirus infections. By analyzing the pharmacophore of the HRV14-

3C enzyme and comparing it with another serotype (HRV2-3C), a better understanding of these sub-genus similarities and differences was achieved. The peptide-based inhibitor used in this study revealed that some inhibitor interactions are essentially conserved across all 3C protease species and the added length of the inhibitor helped rationalize the extended substrate specificity noted with hydrolysis cleavage rate studies using natural 3C protease substrates [40, 39, 44].

The necessity to synthesize the acetyl-LEALFQ-ethyl propenoate inhibitor for the HRV14-3C protease proved pivotal in the success of this structural study. By inactivating the HRV14-3C enzyme, an increase of the protein's stability at NMR concentrations was achieved, such that it remains in solution 4 years after the initial sample prepartion with no change in the ^{15}N-HSQC spectrum. Additionally, an improvement in spectral quality was achieved compared to the apo HRV14-3C enzyme, which allowed for a near complete chemical shift assignment of the protease. The structure and chemical shift data for the inhibited enzyme provided a necessary template to complete the 3D solution structure of the apo enzyme, a task that proved difficult initially due to cluttered spectra resulting from intermediate conformational exchange processes and sparse data for residues involved in substrate interactions. These observations will be expanded upon in the forthcoming chapter that describes the solution structure calculation and characterization of the apo HRV14-3C protease.

Bibliography

[1] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N Chemical Shift Referencing in Biomolecular NMR. *J Biomol NMR*, 6:135–140, 1995.

[2] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: a Multidimensional Spectral Processing System Based on UNIX Pipes. *J Biomol NMR*, 6:277–293, 1995.

[3] B. A. Johnson and R. B. Blevins. NMRView: A Computer Program for the Visualization and Analysis of NMR Data. *J Biomol NMR*, 4:603–614, 1994.

[4] T. C. Bjorndahl, H. Monzavi, and D. S. Wishart. Letter to the Editor: Backbone ^{1}H, ^{15}N and ^{13}C Assignments for the Human Rhinovirus 3C Protease (serotype 14). *J Biomol NMR*, 26:85–86, 2003.

[5] M. Ikura, L. E. Kay, and A. Bax. A Novel Approach for Sequential Assignment of ^{1}H, ^{13}C, and ^{15}N Spectra of Larger Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin. *Biochemistry*, 29:4659–4667, 1990.

[6] M. Wittekind and L. Mueller. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha and Beta-Carbon Resonances in Proteins. *J Magn Reson B*, 101:201–205, 1993.

[7] S. Grzesiek and A. Bax. Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *J Am Chem Soc*, 114:6291–6293, 1992.

[8] D. R. Muhandiram and L. E. Kay. Gradient-Enhanced Triple-Resonance 3-Dimensional NMR Experiments with Improved Sensitivity. *J Magn Reson B*, 103:203–216, 1994.

[9] G. W. Vuister and A. Bax. Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond J(HNHα) Coupling Constants in ^{15}N-Enriched Proteins. *J Am Chem Soc*, 115:7772–7777, 1993.

[10] S. Grzesiek, J. Anglister, and A. Bax. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in $^{13}C/^{15}N$-Enriched Proteins by Isotropic Mixing of ^{13}C Magnetization. *J Magn Reson*, 101:114–119, 1993.

[11] A. Bax, G. M. Clore, and A. M. Gronenborn. ^{1}H-^{1}HCorrelation Via Isotropic Mixing of ^{13}C Magnetization, a New 3-Dimensional Approach for Assigning ^{1}Hand ^{13}CSpectra of ^{13}C-Enriched Proteins. *J Magn Reson*, 88:425–431, 1990.

[12] L. E. Kay, G. Y. Xu, A. U. Singer, D.R. Muhandiram, and J. D. Forman-Kay. A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain ^{1}H and ^{13}C Correlations in H_2O Samples of Proteins. *J Magn Reson B*, 101:333–337, 1993.

[13] L. E. Kay, P. Keifer, and T. Saarinen. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J Am Chem Soc*, 114:10663–10665, 1992.

[14] A. Bax and D.G. Davis. MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. *J Magn Reson*, 65:355–360, 1985.

[15] O. Zhang, L. E. Kay, J. P. Olivier, and J. D. Forman-Kay. Backbone ^{1}H and ^{15}N Resonance Assignments of the N-terminal SH3 Domain of Drk in Folded and Unfolded States Using Enhanced-Sensitivity Pulsed Field Gradient NMR Techniques. *J Biomol NMR*, 4:845–858, 1994.

[16] C. Zwahlen, P. Legault, S. J. F. Vincent, J. Greenblatt, R. Konrat, and L. E. Kay. Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage Lambda N-Peptide/BoxB RNA Complex. *J Am Chem Soc*, 119:6711–6721, 1997.

[17] M. Ikura and A. Bax. Isotope-Filtered 2D NMR of a Protein-Peptide Complex: Study of a Skeletal Muscle Myosin Light Chain Kinase Fragment Bound to Calmodulin. *J Am Chem Soc*, 114:2433–2440, 1992.

[18] G. Cornilescu, F. Delaglio, and A. Bax. Protein Backbone Angle Restraints from Searching a Database for Chemical Shift and Sequence hHomology. *J Biomol NMR*, 13:289–302, 1999.

[19] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, Read R. J., L. M. Rice, T. Simonson, and G. L. Warren. Crystallography & NMR system: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr D Biol Crystallogr*, 54:905–921, 1998.

[20] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, and G. M. Clore. The Xplor-NIH NMR Molecular Structure Determination Package. *J Magn Reson*, 160:65–73, 2003.

[21] A. J. Nederveen, J. F. Doreleijers, W. Vranken, Z. Miller, C. A. E. M. Spronk, S. B. Nabuurs, P. Guentert, M. Livny, J. L. Markley, M. Nilges, Ulrich E. L., R. Kaptein, and A. M. Bonvin. RECOORD: A Recalculated coordinate Database of 500+ Proteins from the PDB Using Restraints from the BioMagResBank. *Proteins*, 59:662–672, 2005.

[22] J.F. Doreleijers, M.L. Raves, J.A.C. Rullmann, and R. Kaptein. Completeness of NOEs in Protein Structure: A Statistical Analysis of NMR. *J Biomol NMR*, 14:123–132, 1999.

[23] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thorton. PROCHECK: A Program to check the Stereochemical Quality of Protein Structures. *J Appl Cryst*, 26:283–290, 1993.

[24] R. W. W. Hooft, G. Vriend, C. Sander, and E. E. Abola. Errors in Protein Structures. *Nature*, 381:272, 1996.

[25] R. Koradi, M. Billeter, and K. Wüthrich. MOLMOL: A Program for Display and analysis of Macromolecular Structures. *J Mol Graphics*, 14:51–55, 1996.

[26] L. Willard, A. Ranjun, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes, and Wishart D. S. VADAR: A Web Server for Quantitative Evaluation of Protein structure Quality. *Nucleic Acids Res*, 31:3316–3319, 2003.

[27] K. Wüthrich. *NMR of Proteins and Nucleic Acids.* Wiley Press, New York, NY, USA, 1986.

[28] G. T. Montelione, K. Wüthrich, E. C. Nice, A. W. Burgess, and H. A. Scheraga. Identification of Two Anti-Parallel Beta-Sheet Conformations in the Solution Structure of Murine Epidermal Growth Factor by Proton Magnetic Resonance. *Proc Natl Acad Sci U S A*, 83:8594–8598, 1986.

[29] A. Pardi, M. Billeter, and K. Wüthrich. Calibration of the angular dependence of the amide proton-Cα proton coupling constants, ^{3}J-HNHα, in a globular protein. Use of ^{3}J-HNHα for identification of helical secondary structure. *J Mol Biol*, 180:741–751, 1984.

[30] D. S. Wishart, B. D. Sykes, and F. M. Richards. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure through NMR Spectroscopy. *Biochemistry*, 31:1647–1651, 1992.

[31] P. Lavigne, J. R. Bagu, R. Boyko, L. Willard, C. F. Holmes, and B. D. Sykes. Structure-Based Thermodynamic Analysis of the Dissociation of Protein Phosphatase-1 Catalytic Subunit and Microcystin-LR Docked Complexes. *Protein Sci*, 9:252–264, 2000.

[32] N Tjandra and A. Bax. Solution NMR Measurement of Amide Proton Chemical Shift Anisotropy in ^{15}N-Enriched Proteins. Correlation with Hydrogen Bond Length. *J Am Chem Soc*, 119:8076–8082, 1997.

[33] J. Yin, E. M. Bergmann, M. M. Cherney, S. M. Lall, R. P. Jain, J. C. Vederas, and M. N. G. James. Dual Modes of Modification of Hepatitis A Virus 3C Protease by a Serine-derived Beta-Lactone: selective crystallization and Formation of a Functional Catalytic Triad in the Active Site. *J Mol Biol*, 354:854–871, 2005.

[34] D. A. Matthews, P. A. Dragovich, S. A. Webber, S. A. Fuhrman, A. K. Patick, L. S. Zalman, T. F. Hendrickson, R. A. Love, T. J. Prins, J. T. Marakovits, R. Zhou, J. Tikhe., C. E. Ford, J. W. Meador III, R. A. Ferre, E. L. Brown, S. L. Binford, M. A. Brothers, D. M. DeLisle, and S. T. Worland. Structure-Assisted Design of Mechanism-Based Irreversible Inhibitors of Human Rhinovirus 3C Protease with Potent Antiviral Activity Against Multiple Rhinovirus Serotypes. *Proc Natl Acad Sci U S A*, 96:11000–11007, 1999.

[35] S. C. Mosimann, M. M. Cherney, S. Sia, S. Plotch, and M. N. G. James. Refined X-ray Crystallographic Structure of the Poliovirus 3C Gene Product. *J Mol Biol*, 273:1032–1047, 1997.

[36] D. A. Matthews, W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, C. L. Gribskov, and Worland S. Structure of Human Rhinovirus 3C Protease Reveals a Trypsin-like Polypeptide Fold, RNA-Binding Site, and Means for Cleaving Precursor Polyprotein. *Cell*, 77:761–771, 1994.

[37] P. Hudaky and A. Perczel. Toward Direct Determination of Conformations of Protein Building Units from Multidimensional NMR Experiments VI: Chemical Shift analysis of His to Gain 3D Structure and Protonation State Information. *J Comp Chem*, 26:1307–1317, 2005.

[38] A. C. Wallace, R. A. Laskowski, and Thornton J. M. LIGPLOT: A Program to Generate Schematic diagrams of Protein-Ligand Interactions. *Protein Eng*, 8:127–134, 1995.

[39] M. G. Cordingley, R. B. Register, P. L. Callahan, V. M. Garskey, and R. J. Colonno. Cleavage of Small Peptides In Vitro by Human Rhinovirus 14 3C Protease Expressed in *Escherichia coli*. *J Virol*, 63:5037–5045, 1989.

[40] D. C. Orr, A. C. Long, J. Kay, B. M. Dunn, and J. M. Cameron. Hydrolysis of a Series of Synthetic Peptide Substrates by the Human Rhinovirus 14 3C Proteinase, Cloned and Expressed in *Escherichia coli*. *J Gen Virol*, 70:2931–2942, 1989.

[41] T. O. Johnson, Y. Hua, H. T. Luu, E. L. Brown, F. Chan, S. S. Chu, P. S. Dragovich, B. W. Eastman, R. A. Ferre, S. A. Fuhrman, T. F. Hendrickson, F. C. Maldonado, D. A. Matthews, J. W. Meador III, A. K. Patick, S. H. Reich, D. J. Skalitzky, S. T. Worland, M. Yang, and L. S. Zalman. Structure-Based Design of a Parallel Synthetic Array Directed Toward the Discovery of Irreversible Inhibitors of Human Rhinovirus 3C Protease. *J Med Chem*, 45:2016–2023, 2002.

[42] P. S. Dragovich, T. J. Prins, R. Zhou, S. A. Fuhrman, A. K. Patick, D. A. Matthews, C. E. Ford, J. W. Meador III, R. A. Ferre, and S. T. Worland. Structure-Based Design, Synthesis, and Biological Evaluation of Irreversible Human Rhinovirus 3C Protease Inhibitors. 3. Structure-Activity Studies of Ketomethylene-Containing Peptidomimetics. *J Med Chem*, 42:1203–1212, 1999.

[43] S. L. Binford, M. A. Maldonado, P. T. Brothers, L. S. Weady, J. W. Zalman, D. A. Meador III, and A. K. Patick. Conservation of Amino Acids in Human Rhinovirus 3C Protease Correlates with Broad-Spectrum Antiviral Activity of Rupintrivir, a Novel Human Rhinovirus 3C Protease Inhibitor. *Antimicrob Agents and Chemo*, 49:619–626, 2005.

[44] Q. M. Wang, R. B. Johnson, G. A. Cox, E. C. Villarreal, and R. J. Loncharich. A Continuous Colorimetric Assay for Rhinovirus-14 3C Protease Using Peptide p-Nitroanilides as Substrates. *Anal Biochem*, 252:238–245, 1997.

[45] A. C. Long, D. C. Orr, J. M. Cameron, B. M. Dunn, and J. Kay. A Consensus Sequence for Substrate Hydrolysis by Rhinovirus 3C Protease. *FEBS L*, 258:75–78, 1989.

[46] G. A. Cox, R. B. Johnson, J. A. Cook, M. Wakulchik, M. G. Johnson, E. C. Villarreal, and Q. M. Wang. Identification and Characterization of Human Rhinovirus-14 3C Protease Deamidation Isoform. *J Biol Chem*, 274:13211–13216, 1999.

[47] S. H. Chen, J. Lamar, F. Victor, N. Snyder, R. Johnson, B. A. Heinz, M. Wakulchik, and Q. M. Wang. Synthesis and Evaluation of Tripeptidyl Alpha-Ketoamides as Human Rhinovirus 3C Protease Inhibitors. *Bioorg and Med Chem Letters*, 13:3531–3536, 2003.

[48] D. A. Jewel, W. Swietnicki, B. M. Dunn, and B. A. Malcolm. Hepatitis A Virus 3C Proteinase Substrate Specificity. *Biochemistry*, 31:7862–7869, 1992.

Chapter 5

Chemical Shift Assignment and Structure Calculation of the Apo HRV14-3C Protease

5.1 Introduction

The chemical shift and structural data for the inhibited enzyme provided the necessary template to complete the chemical shift assignments and structure calculation for the apo enzyme. This chapter outlines these tasks by presenting both the NMR experiments and structure calculation strategy used. A geometric evaluation of the resulting structures is presented and compared with the structures of the inhibited enzyme. Furthermore, the observed stability differences between the apo and bound states incited the investigation of the enzyme's slow time scale dynamics with $^{1}H/^{2}H$ exchange. These experiments and their results are discussed along with insights into the possible allosteric communication mechanism between the proteolytic and RNA binding interfaces that has been shown to exist for a homologous 3C protease.

5.2 Methods and Materials

5.2.1 NMR Sample Preparation

The HRV14-3C protease was expressed and purified using the methods outlined in *Appendix B* and discussed in detail in Chapter 2. [U-^{15}N]-labelled and [U-^{13}C/^{15}N]-labelled samples were prepared by substituting ^{15}N-labelled NH_4Cl and/or ^{13}C-labelled glucose for

the unlabelled counterparts in the MM recipe outlined in *Appendix A*. NMR samples were prepared by dialyzing the purified protease into 20 mM KH_2PO_4 buffer (pH 6.5, 0.5 mM EDTA, 15 mM DTT) and concentrated to ~0.5 - 1 mM *via* ultrafiltration. Ten percent D_2O was added for maintaining the spectrum lock and 0.1 mM DSS was added for internal referencing [1]. Samples were filtered through 22 μm epindorph filters and placed in either 5 mm thin-walled WILMAD™ NMR tubes or 3mm SHIGEMI™ NMR tubes. Samples were recovered between the collection of backbone assignment NMR experiments by using the refolding protocol outlined in section 2.2.8. Production of the apo protease subsequent to the collection of the backbone assignment experiments (Table 5.1) employed a hypoxyapatite column (*Appendix B*) and the addition of 0.1 mM NaN_3. To determine D_2O exchange rates, samples of the apo and inhibited enzyme were exchanged into 99.9% D_2O containing 20 mM KH_2PO_4, 15 mM DTT, 0.5 mM EDTA at pH 6.5. This resulted with a final D_2O concentration of 99.6% and a corrected pD of 6.9. The exchange was performed by 6 successive volumentric dilutions and subsequent concentration *via* ultracentrifugation and was done at 25 °C.

5.2.2 NMR Data Collection

All experiments conducted on the apo HRV14-3C protease were done at 25 °C using Varian 500, 600 and 800 MHz INOVA spectrometers. The 500 MHz spectrometer was fitted with either a 5 mm HCN z-gradient PFG room temperature probe or a Z-gradient PFG Varian coldprobe. The 600 MHz spectrometer was fitted with a 5 mm HCN z-gradient PFG room temperature probe. The 800 MHz spectrometer was equipped with a 5mm HCN xyz-gradient PFG coldprobe. All experiments were conducted using Varian Protein Pack pulse sequences (VNMR v3.1c or VNMRJ).

The 2D ^{15}N-HSQC [2] experiment was collected on a [U-^{15}N]-labelled sample (Figure 5.1). HNCA [3], HNCACB [4], CBCA(CO)NH [5], HNCO [6] and HNHA [7] spectra were collected on [U-^{13}C/^{15}N]-labelled samples. Spectra were processed with NMRPIPE [8] and further analyzed with NMRVIEW [9]. Hydrogen exchange data for the apo HRV14-3C protease was acquired with a series of two-dimensional ^{15}N-HSQC spectra collected at 0, 38, 99, 172, 486 and 6580 minutes following H_2O to D_2O exchange of the ^{13}C/^{15}N-labelled sample. For the inhibited HRV14-3C protease, hydrogen exchange data were acquired from

a series of two-dimensional ^{15}N-HSQC spectra collected on the [U-^{13}C/^{15}N]-labelled sample at 0, 156, 312, 1220, 2765 and 8702 minutes following H_2O/D_2O solvent exchange. The exchange experiments were conducted at 25 °C. To acquire T_2 measurements, one and two dimensional ^{15}N CPMG-HSQC spectra of the amide signals were collected on the apo and inhibited [U-^{13}C/^{15}N]-labelled samples using delays of 10, 30, 50, 70 and 90 ms.

Table 5.1: **Backone NMR Experiments for the Apo HRV14-3C Protease**

	Nucleus			Number of Points			Spectral Width (Hz)			
Experiment	t1	t2	t3	t1	t2	t3	t1	t2	t3	Transients
Backbone Assignment Experiments										
HSQC	^{15}N	^{1}H		512	928		2000	6000		40
HNCO	^{13}C	^{15}N	^{1}H	64	32	896	3018	2000	6000	24
HNHA†**	^{1}H	^{15}N	^{1}H	64	20	1024	6000	2000	6000	40
HNCA	^{13}C	^{15}N	^{1}H	70	24	896	3770	2000	6000	32
HNCACB	^{13}C	^{15}N	^{1}H	64	32	896	10054	1800	6982	32
CBCA(CO)NNH†	^{13}C	^{15}N	^{1}H	50	32	894	9000	1700	6982	32
Sidechain Assignments										
C(CO)NNH†	^{13}C	^{15}N	^{1}H	64	32	768	10054	2000	6000	32
H(C CO)NNH†	^{1}H	^{15}N	^{1}H	58	32	1024	6000	2000	6000	32
HCCH-TOCSY‡ ††	^{1}H	^{13}C	^{1}H	128	32	1024	8000	10056	8000	16
^{13}C HSQC (35 ppm§)	^{13}C	^{1}H		256	1024		12568	6000		48
Restraint Assignments										
^{15}N-NOESY-HSQC (τ_m 80 ms ¶)	^{1}H	^{15}N	^{1}H	64	32	1024	6000	2000	6000	32
^{13}C-NOESY-HSQC (35 ppm§, τ_m 50 ms¶)†	^{1}H	^{13}C	^{1}H	128	34	128	6000	10056	6000	16
^{13}C-NOESY-HSQC (35 ppm§, τ_m 100 ms¶)‡‖	^{1}H	^{13}C	^{1}H	128	32	1534	6000	10000	11990	16
D_2O exchange ^{15}N HSQCs‡*	^{15}N	^{1}H		32	1024		2128	8012		32

† Conducted at 500MHz with a cold probe

** Sample concentration ~0.5 mM

†† Conducted at NANUC at 500MHz

§ Carrier frequency

¶ Mixing time

‖ Conducted at NANUC at 800MHz

‡ Conducted on a sample in 99.6% D_2O

* Conducted at NANUC at 600MHz. The first spectrum in the series had 16 transients, the fourth spectrum in the series (collected 448 at minutes post exchange) had 256 increments in the first indirectly detected dimension

A protein sample in 90% H_2O/10% D_2O was required for the T_2 measurements of the apo enzyme, while the residual amide envelope following D_2O exchange was sufficient to collect data for the inhibited HRV14-3C protease (Table 4.1). The integration range for both sets of data was between 6.0 and 11.0 ppm. Relaxation delays of 2 seconds were used for all ^{15}N T_2 measurements. The experimental parameters used to obtain chemical shift assignments and structural restraint data for the apo HRV14-3C protease is provided in Table 5.1.

5.2.3 Correlation Time Calculation

The protein correlation time (τ_c) was calculated from the formula 5.1 [10]:

$$\frac{1}{T_{2,\ HN}} = 1.11 \cdot \tau_c \tag{5.1}$$

which is derived from the equation:

$$\frac{1}{T_2} = \frac{9}{20} \cdot \frac{\gamma^4 \hbar^2}{r^6} \cdot \tau_c \tag{5.2}$$

where γ is the detected nucleus' gyromagnetic ratio, $\hbar$ is Plank's constant $\div\ 2\pi$ or $\frac{h}{2\pi}$ (1.055^{-34} J . s) and r is the internuclear distance in Angstroms.

The $T_{2,HN}$ was calculated by fitting the integrated ^{15}N CPMG-HSQC 1D amide envelope signal intensity to the exponential curve defined by:

$$I = I_o \cdot e^{-kt} \tag{5.3}$$

VNMR v6.1c was used to integrate the 1D amide signal and the ZUNZUN web server (`http://zunzun.com/`) was used to fit the data. Estimation of the correlation time was calculated using the Debye-Stokes-Einstein law:

$$\tau_c = \frac{\eta 4\pi r_H^3}{3k_B T} \tag{5.4}$$

where T was 298 °K, k_B is Boltzmann's constant (1.3806503 x 10^{-23} m^2 kg s^{-2} K^{-1}) and η is the viscosity of water (0.89 cp at 25 °C). The r_H (the hydrodynamic radius) was determined to be 21.69 Å using the formula [11]:

$$r_H = \sqrt[3]{\frac{3\bar{V}M_r}{4\eta N_A}} + r_W \tag{5.5}$$

where r_W is the radius of a water molecule at 1.6 Å, $\bar{V}$ is the enzyme's specific volume (estimated at 0.73cm^3/g), N_A is Avogadro's number and M_r is the mass of the isotopically labelled protein (21,838.66 g/mol as determined by MALDI-TOF mass spectrometry).

Calculation of the fractional monomeric form was done using the following equation:

$$\tau_c = \tau_{c,M} \cdot [f_M + 2 \cdot (1 - f_M)] \tag{5.6}$$

Where τ_c is the calculated corrlation time. $\tau_{c,M}$ is the predicted monomeric correlation time derived from equation 5.4 and f_M is the fraction of monomeric protein.

5.2.4 Hydrogen/Deuterium Exchange Rate Analysis

All $^1H/^2H$ ^{15}N-HSQC exchange spectra were processed with identical phasing and apodization functions. Peaks were auto-picked and volumes calculated with NMRVIEW [9]. The peak volumes vs. exchange time were fit to the first order exponential decay equation:

$$V = V_o \cdot e^{-kt} + baseline \tag{5.7}$$

using a non-linear least squares fit routine in the ORIGIN (v7.5) software package to obtain K_{ex} rates. Protection factors (P_{factor}) for each residue were calculated by comparing the calculated exchange rate (K_{ex}) with the predicted random coil exchange rates (K_{rc}) where:

$$P_{factor} = K_{rc}/K_{ex} \tag{5.8}$$

The random coil rates were calculated using the methods described by Bai *et al.* [12] and corrected for temperature and pD differences [13] prior to P_{factor} calculation.

5.3 Results

5.3.1 Apo HRV14-3C Chemical Shift Assignments

The experimental methodology used to assign the inhibited HRV14-3C protease (discussed is Chapter 4: Section 4.3.1) was adopted for completing the backbone chemical shift assignments of the apo HRV14-3C protease. The inhibited enzyme's chemical shifts and structure were used to guide the assignment process. Aromatic assignments were obtained from NOESY data during the apo enzyme's structure calculation. The apo samples pre-

pared for collecting data subsequent to the backbone assignment experiments were ~25% more concentrated compared with the inhibited HRV14-3C samples. This fact, in addition to using a coldprobe, greatly improved experimental sensitivty and overall spectral quality. This is illustrated by comparing the spectra for the apo and inhibited enzymes. The additional side-chain chemical shift data obtained was added to the initial deposited backbone chemical shifts (BMRB # 5659) along with a few corrections to the original backbone assignment data. These changes primarily involved regions of the protein that displayed broadened chemical shift peaks.

92% of all possible backbone shifts were obtained, which included 85% of $^{1}H_{\alpha}$, 83% of $^{13}C'$, 90% of $^{1}H_{N}$, 88% of ^{15}N, 98% of $^{13}C_{\alpha}$ and 97% of $^{13}C_{\beta}$ chemical shifts. In all, 1817 of 2171 possible ^{13}C, ^{15}N and ^{1}H assignments were obtained for 180 of 182 residues in apo HRV14-3C protease (84% complete). These chemical shifts are presented in *Appendix D.*

5.3.2 Apo NMR HRV14-3C Structure Calculation

Preliminary structures of the apo HRV14-3C enzyme were created with CYANA [14] and the inhibited HRV14-3C protease (PDB code 2B0F) was used as a starting structure. ^{15}N and ^{13}C-edited NOESY peak lists were manually assigned using NMRVIEW [9] and used to generate starting NOE restraint lists. These lists were used during the structure calculation with CYANA and were subsequently refined by CANDID [15]. Twenty-one spurious NOE peaks were identified using the NOAH algorithm [16] during the calculations and removed. The refined peak lists were then fed back into NMRVIEW for manual confirmation. These '*manicured*' peak lists provided 1515 non-redundant NOE assignments (533 short range, 535 medium range, 447 long range) from the ^{15}N-NOESY-HSQC and ^{13}C-NOESY-HSQC spectra [17]. The final set of NOEs were calibrated using proton cross peak intensities and binned with upper bounds of 3.0, 4.2 and 5.5 Å corresponding to strong, medium and weak intensities respectively. All lower bounds were set to 1.8 Å. Additionally, 29 stereospecific assignments were made with HABAS during the CYANA simulated annealing structure calculations and incorporated into subsequent structure calculations. $^{3}J_{HNH\alpha}$ coupling constants, 129 in total, were unambiguously determined from the HNHA spectrum and used to assign 129 phi (ϕ) angles. One-hundred psi (ψ) angles, which clustered in favorable regions of the Ramachandran plot following the initial structure calculation, and matched

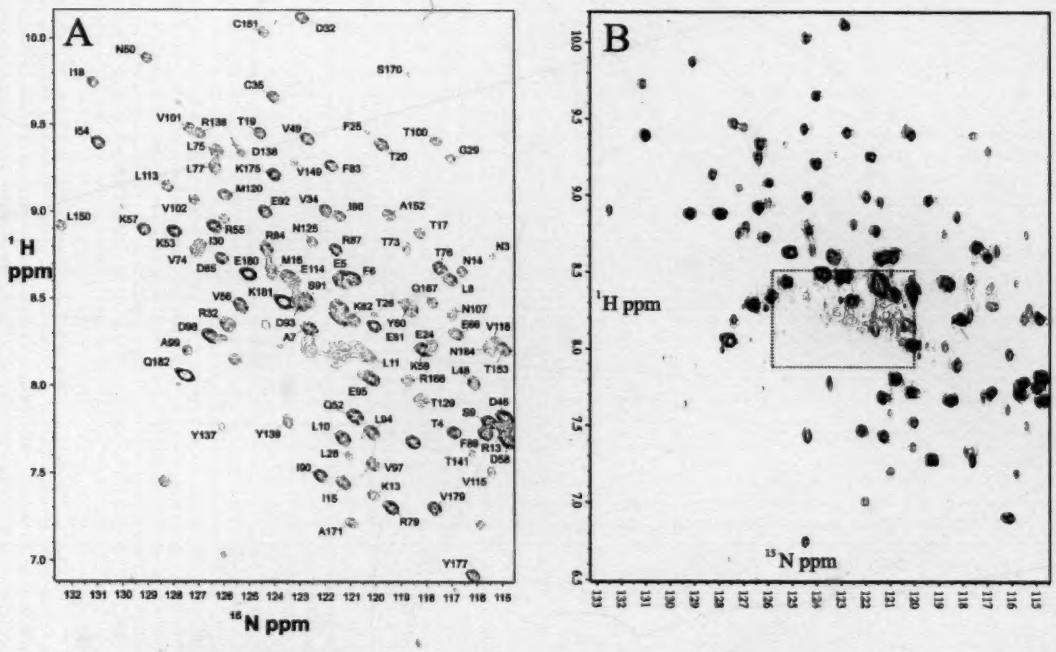

Figure 5.1: **^{15}N-HSQC Spectrum of the Apo HRV14-3C Protease.** **A.** A portion of the ^{15}N-HSQC spectrum for the HRV14-3C protease is shown. Labeled peaks correlate with assignments listed in Table D.3. **B.** An overlay of two ^{15}N-HSQC spectra collected on different apo HRV14-3C samples prepared by different methods (RED: no hypoxyapatite column, sample stability ~1 week. BLUE: hypoxyapatite column purification and the addition of 0.1 mM NaN_3, sample stability >6 months). The spurious peaks visible in the '*random coil*' region of the spectra (ν_N ~122; ν_{HN} ~8.25 ppm) were initially thought to arise from degradation as the sample precipitated. However, spectra collected on the sample prepared with improved stability also displayed these additional peaks, which are now believed to arise from conformational exchange processes.

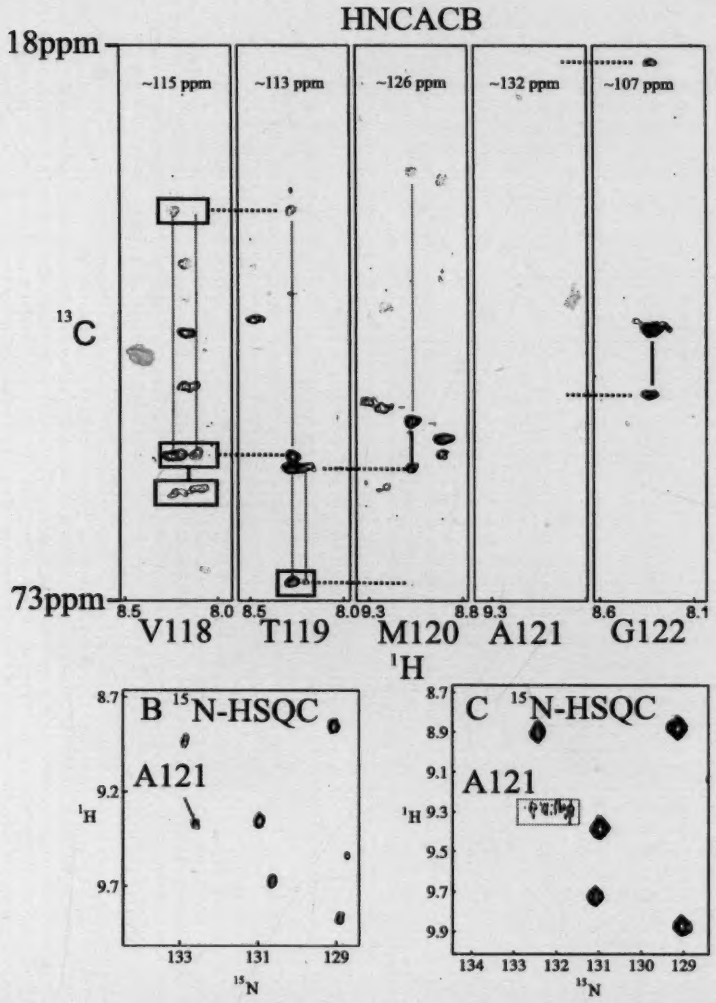

Figure 5.2: **HNCACB Strip Plots from the Apo HRV14-3C Protease**. A walkthough of a portion of the apo HRV14-3C HNCACB spectra. $^{13}C_{\beta}$ peaks are colored red. Broken lines connect the *i* with *i-1* chemical shifts for consecutive strip plots. Slowly exchanging residues presenting multiple amide peaks are boxed. The chemical shift information for Ala^{121} was not attainable from this experiment. Comparison of the ^{15}N-HSQC spectra for the inhibited HRV14-3C protease (**Spectrum B**) with apo HRV14-3C protease (**Spectrum C**) shows extensive peak broadening due to chemical exchange in the apo enzyme. These spectra were collected at 500 MHz using a Varian INOVA spectrometer equipped with a room temperature HCN probe. Spectrum C is shown at a much lower contour level in order to view the broadened NMR peak of Ala^{121}.

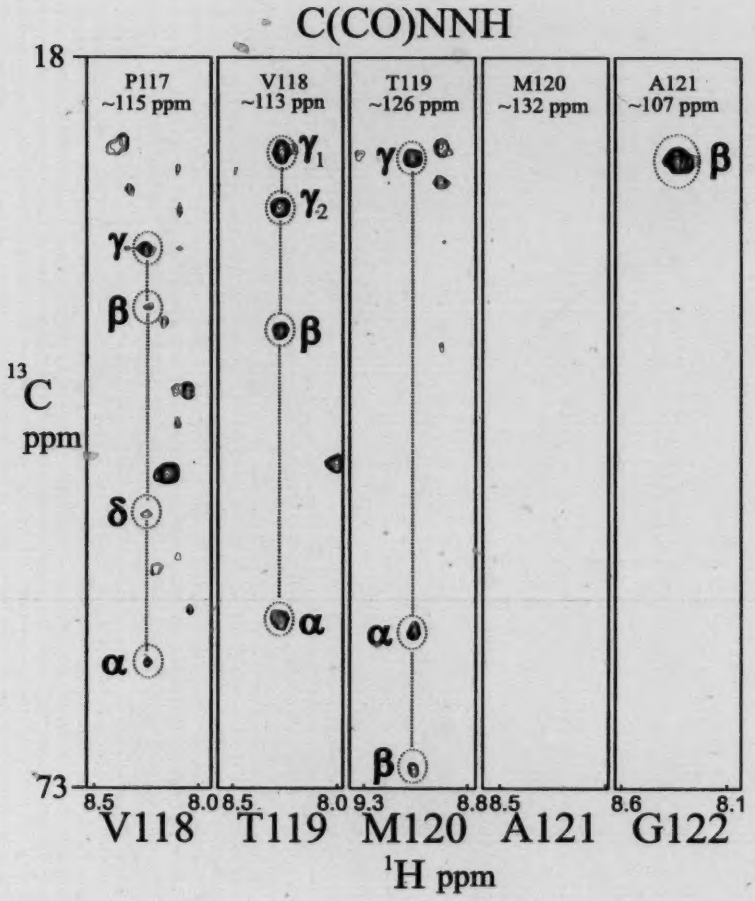

Figure 5.3: **C(CO)NNH Strip Plots from the Apo HRV14-3C Protease.** Strips corresponding to the amide planes for residues Val^{118} to Gly^{122} are shown. The preceding residue labels are inserted at the bottom of each strip plot. Broken lines connect intraresidue ^{13}C peaks. This spectrum was collected at 500 MHz using a Varian INOVA spectrometer fitted with a coldprobe.

TALOS [18] and SHIFTOR (`http://redpoll.pharmacy.ualberta.ca/shiftor`) predictions, were also assigned. Limits of ±40° were applied to the dihedral restraints. These limits were inflated to ±60° for residues Asn^{64}, Glu^{92}, Leu^{94}, Val^{97}, Ser^{105}, Thr^{141}, Gly^{144} and Lys^{174} to account for weak signals in the HNHA spectrum and subsequent increased errors in the measured $^{3}J_{HNH\alpha}$ data. Like the structure calculations for the inhibited enzyme, ϕ and ψ angle restraints were assigned following the initial structure calculations with NOE data alone. The χ angles for His^{40} and Glu^{71} were restrained to values that matched both the inhibited HRV14-3C protease (PDB code 2B0F) and the previously reported apo HRV14-3C protease structure [19] in order to maintain the orientation of the active-site triad residues. Hydrogen bonds were identified following analysis of the ^{15}N-HSQC spectra collected on the [U-^{13}C/^{15}N]-HRV14-3C protease sample exchanged into 99.6% D_2O buffer. A total of 64 hydrogen bond donors were assigned to amides presenting signals after 38 minutes following deuterium exchange at 25 °C, pD 6.9 and assigned limits of 1.5 - 2.5 Å for H_N to O distances and 2.5 - 3.5 Å for N to O distances. Center-weighted pseudo-atom corrections were used for ambiguous methylene and methyl proton NOE distances. 300 initial structures were calculated *via* the simulated annealing protocol (high temperature steps = 24,000 at 1000 °K; cooling steps = 12,000; final temperature = 100 °K) using CNS v.1.1 [20]. Fifty of the lowest energy structures with were refined in explicit solvent using the RECOORD protocol [21]. The PARALLHDG and OPLSX non-bonded parameter sets were employed for the simulated annealing and water refinement protocols respectively. The top 20 structures, which had no NOE violations >0.2 Å and no dihedral

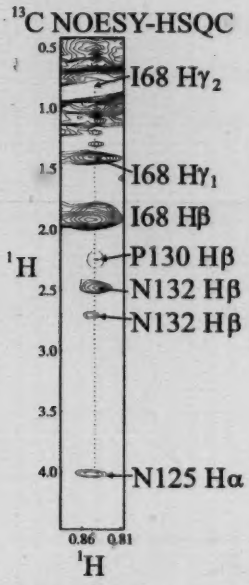

Figure 5.4: ^{13}C -NOESY-HSQC Spectrum for Ile^{68} $^{1}H_{\gamma 2}$ of the Apo HRV14-3C Protease. NOE data for the $^{1}H_{\gamma 2}$ of Ile^{68} shows cross domain NOEs to Asn^{125} $^{1}H_{\alpha}$, Pro^{130} $^{1}H_{\beta}$ and Asn^{132} $^{1}H_{\beta}$ atoms.

angle violations >5°, have been deposited into the BMRB under accession code 2IN2.

An average of 8.5 NOEs per residue were used for the structure calculations. 2.5 long-range NOEs per residue and 3.5 medium-range NOE restraints per residue yielded structures with good structural statistics and precision (Table 5.2).

Table 5.2: **Structural Statistics for the HRV14-3C Proteases***

	Inhibited	Apo
Distance Restraints		
All NOE distances	1998	1515
Intra-residue (protein)	1911	1515
Sequential (| i-j | = 1)	749	533
Medium (1 <| i-j |≤ 4)	690	535
Long (| i-j |> 4)	472	447
Inter-residue (protein-inhibitor)	76	
Intra-residue (inhibitor)	11	
Hydrogen bonds	87	64
Violations		
Structures with violations > 0.3Å	0	0
Structures with violations > 0.2Å	6	0
Dihedral angle restraints		
All	258	231
ϕ	131	129
ψ	126	100
χ_1	1	3
Ramachandran Plot†		
Residues in most favored region	78.90%	79.40%
Residues in additionally allowed region	19.90%	19.40%
Residues in generously allowed region	0.50%	0.60%
Residues in disallowed region	0.60%	0.60%
WHAT-CHECK scores‡		
Second generation packing	-1.25	-0.53
χ_1 / χ_2	-1.74	-0.89
RMSD to mean structure§		
Backbone	0.82 ± 0.13	1.07 ± 0.17
Heavy atom	1.49 ± 0.20	1.56 ± 0.25
Region 15-78		
Backbone	0.72 ± 0.14	0.87 ± 0.22
Heavy atom	1.41 ± 0.20	1.64 ± 0.31
Region 99-103, 111-172		
Backbone	0.56 ± 0.12	1.00 ± 0.21
Heavy atom	1.17 ± 0.17	1.57 ± 0.25

*Data for both forms of the HRV14-3C protease are presented for comparison
†Calculated with PROCHECK-NMR [22]
‡Calculated with WHAT-CHECK [23]
§Calculated with MOLMOL [24]

None of the deposited structures had violations greater than 0.2 Å. The total number of NOEs used in the structure calculation was lower than that used for the inhibited HRV14-3C structure calculation. However, the resulting structures were statistically comparable. The majority of NOEs not used were sequential, which provide minimal contribrution to the

overall structure calculation and global fold. It should be noted however, that the initial number of structures generated was increased 3-fold compared to the inhibited HRV-3C protease structure calculations in order to acquire an ensemble with minimal violations.

Compared with the inhibited enzyme, the apo form of the HRV14-3C protease presented larger RMSDs in areas of the protein that are involved with substrate binding (Figure 5.6). The larger RMSDs reported in these regions result from a breakdown of J-coupled connectivity and a subsequent reduction in NOE data. Although, NOE data was identified that localized the $\mathbf{II}_d$ - $\mathbf{II}_e$ β-strands into a conformation similar to the inhibited HRV14-3C, the inhibited HRV2-3C and apo Polio-3C enzymes. The significance of this is that early structures for the homologous HAV-3C protease showed large amplitude correlated motions for these β-strands. The cross domain NOE-data that helped localize these strands in the apo HRV14-3C enzyme are shown in Figure 5.4. These data suggest that large amplitude correlated motions do not occur between the two domains in the apo HRV14-3C protease.

Geometric and structure quality analysis for both structures was carried out using VADAR [25]. AQUA [26] and PROCHECK-NMR [22] were used to calculate and analyze NOE and dihedral angle restraint violations. The final structure ensemble presented good statistical results following RECOORD water refinement [21] with What-Check packing and rotamer Z scores [23] of -0.53 and -0.89 respectively. Equivalent X-ray resolutions based on Ramachandran plot quality assessment (2.3 Å) and χ_1 pooled and χ_2 trans angle standard deviation assessments (1.0 Å) were calculated with PROCECK-NMR [22]. These results are comparable to the inhibited structure (2.4 Å and 1.1 Å) which confirms that the structures are of comparable quality.

5.4 Discussion

5.4.1 Apo vs. Inhibited HRV14-3C Protease Chemical Shift Assignments

There were an increased number of un-assignable amide backbone chemical shifts (9%) for the apo enzyme in comparison to the inhibited HRV14-3C (section 4.3.1). In addition to these ^1H_N peaks, a number of broad NMR peaks were visible. It is believed that this linewidth heterogeneity results from intermediate (μs) conformational exchange processes. Examples of these broad NMR signals for the HRV14-3C protease are shown in the

^{15}N-HSCQ spectra in Figures 5.1 and 5.2. The C(CO)NNH spectrum (Figure 5.3) shows a markedly reduced peak intensity for the peaks correlating back to the $^{1}H_{N}$ of Val^{118}. These two spectra were collected on different samples (one purified using the hypoxyapatite column) and at different time periods after sample preparation. Furthermore, both sets of experiments display similar findings for the $^{1}H_{N}$ of Ala^{121}. These facts ruled out the possibility that these additional peaks resulted from degradation.

Broader amide NMR signals were assigned to residues residing in several different regions of the protein. One region involved the proteolytic interface, which also had a number of un-assignable amide signals (Asn^{105}, Asn^{110}, Leu^{123}, Ser^{127}, Ser^{128}, Thr^{131}, Arg^{133} - Ile^{135}, Lys^{142}, and Gln^{145}). On the opposite side of the protein are the RNA binding [27] and the 3CD hetero-dimerization [28] interfaces that involved the residues $KFRDI^{82-86}$ and DLE^{93-95} respectively. The residues that presented un-assignable or very broad NMR signals are mapped onto a ribbon representation of the protein in Figure 5.5. There is a strong correlation between the structural proximity of these residues and their association with biological processes. The remaining residues presenting broad NMR peaks (Val^{56}, Lys^{61}, Leu^{62}, Asn^{67}, Thr^{76} and Asn^{80}) appear to localize linearly on a common face of the enzyme. However, these have not yet been associated with any biological function.

Interstingly, Val^{56} resides in a surface exposed β-strand that was implicated in a homo-dimerization interface in the homologous apo Polio-3C protease's X-ray structure [29]. However, besides this residue, no other residue within this β-strand (Figure 4.7: $\mathbf{I}_e$) for the apo HRV14-3C protease presented broadened NMR peaks.

5.4.2 Apo vs. Inhibited HRV14-3C Protease Solution Structure

An increase of the global RMSD relative to the inhibited enzyme was observed for the apo enzyme, which resulted from the breakdown of J-coupled connectivity data and the subsequent limited amount of NOE data. Despite the improved NMR sample stability of the apo enzyme afforded with the hydroxyapatite column purification and addition of 0.1 mM NaN_3, the connectivity information remained sparse in the areas of the protein that are involved with proteolytic substrate interaction. These data are in contrast to the data for the inhibited enzyme, which presented well-resolved NMR spectra and allowed near complete chemical shift assignments. The apo enzyme RMSDs were calculated to be 1.07 ±

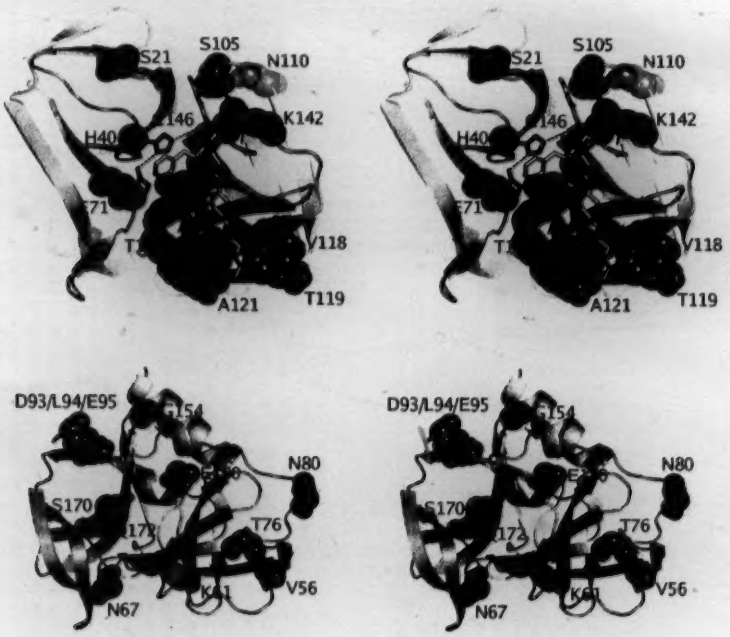

Figure 5.5: **Suspected Residues Undergoing μs Conformational Exchange**. Two orientations of the apo HRV14-3C protease are presented to show the majority of residues undergoing chemical exchange. The top orientation shows the proteolytic site and substrate binding region. The majority of un-assignable amide signals are located within this region. The bottom orientation is the opposite side of the protease, in which the RNA binding interface resides. Residues with unassigned signals are colored red and residues with broad NMR signals are colored blue. It is suspected that these residues undergo μs time scale conformational exchange.

0.17 Å and 1.56 ± 0.25 Å for backbone heavy and all heavy atoms respectively (region: Gly^{1} to Glu^{180}). These increased values relative to the inhibited enzyme (0.82 ± 0.13 and 1.49 ± 0.20) resulted primarily from the flexible, substrate binding C-terminal domain (Ala^{99} to Glu^{180}) that alone yielded RMSDs of 1.00 ± 0.21 Å and 1.57 ± 0.25 Å for backbone and all heavy atoms respectively.

The N-terminal β-barrel domain yielded RMSDs of 0.88 ± 0.23 Å and 1.64 ± 0.31 Å for all heavy atoms from residues Ile^{15} to Arg^{79}. A graphical representation of the atomic displacement data is presented in Figure 5.6. The calculated atomic displacements were normalized using methods described by Billeter *et al.* [30, 31]. The results shown here for

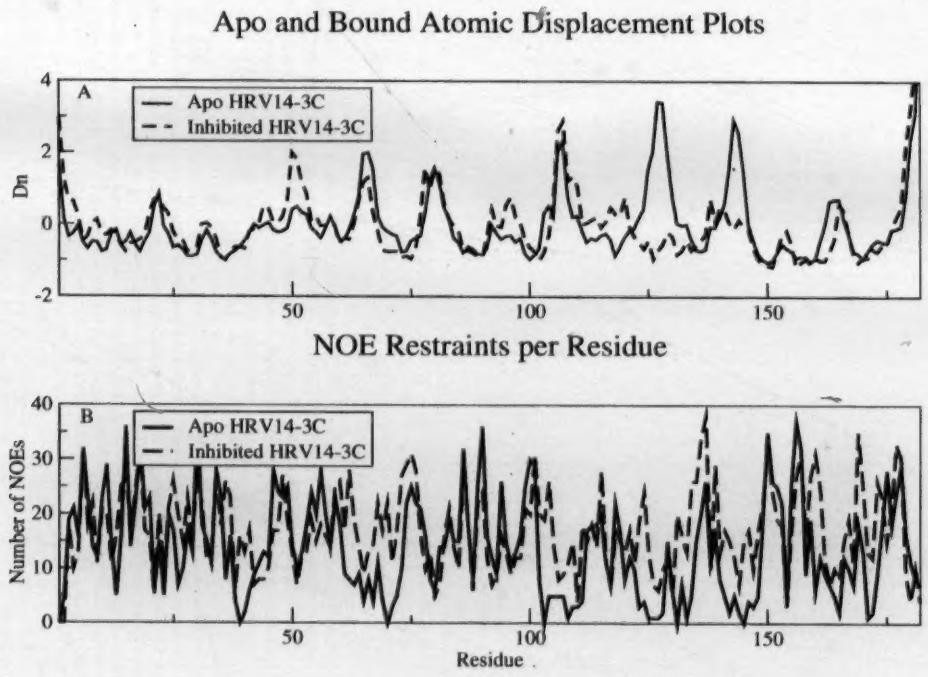

Figure 5.6: **Normalized Atomic Displacements and Per-residue NOE Restraint Count for the Apo and Inhibited HRV14-3C Proteases. A**. The mean atomic displacements (D_n) for the apo and inhibited HRV14-3C enzymes (2IN2.pdb and 2BOF.pdb respectively) are shown. The significant differences between the two enzyme states are localized in areas involved with inhibitor/substrate binding. **B**. The NOE restraints per residue indicate that the areas of the protein structure that exhibit larger atomic displacements correlate with fewer NOE restraints per residue.

the apo HRV14-3C protease are strikingly similar to the normalized B-factors shown for the apo Polio-3C protease (Figure 3.1). The larger normalized atomic displacements and B-factors observed for the atoms involved with substrate binding within the C-terminal domain for both apo 3C proteases (HRV14-3C and Polio) indicates that this domain is inherently flexible and suggests this flexibility is required to bind the various natural peptide substrates. Furthermore, the sparse NMR data available for this region (discussed in section 5.4.1 and shown in Figure 5.6) suggests that μs time scale motions occur.

In contrast to the conformational flexibility observed within the C-terminal domain, which binds the peptide substrate (Figure 5.6), the RNA binding site remains superposable

in either the active and the inhibited states. This finding is indicated by their superposable backbone atomic positions (Table 5.3), which are 0.28 Å.

Table 5.3: **Superposition of Picornaviridae $3C^{pro}$ RNA Binding Region**

Protease	PDB code	Superposed RNA region	Pairwise fit (Å)†
HRV14-3C§‡	2IN2	HD^{31-32} $KFRDI^{82-86}$ I^{156}	0.28
HRV2-3C‡	1CQQ	YD^{31-32} $KFRDI^{82-86}$ I^{157}	0.51
POLIO-3C	1L1N	HD^{31-32} $KFRDI^{82-86}$ I^{157}	0.63 ± 0.13
HAV-3C	1HAV	KD^{35-36} $KFRDI^{95-99}$ I^{187}	0.57 ± 0.06
HAV-3C‡	1QA7	KD^{35-36} $KFRDI^{95-99}$ I^{187}	0.61 ± 0.03
HAV-3C‡	2AO4	KD^{35-36} $KFRDI^{95-99}$ I^{187}	0.51
HAV-3C‡	2CXV	KD^{35-36} $KFRDI^{95-99}$ I^{187}	0.29
FMD-3C	2BHG	FG^{37-38} $KVRDI^{95-99}$ I^{177}	0.44 ± 0.40

†Backbone heavy atom pairwise fits to 2B0F.pdb. Calculationed with MOLMOL [24]
§Comparison between mean NMR structures
‡Structures with bound inhibitors

This region also presents some well-structured backbone hydrogen bonds that are formed between the RNA motif, $KFRDI^{82-86}$, residue Asn^{14} and the loop region His^{31} - Val^{34} in the inhibited state (Figure 5.7). Asn^{14} is a conserved residue for all picornaviridae 3C proteases and is located at a junction between the N-terminus α-helix, **A** (Figure 4.7), and the first β-strand, $\mathbf{I}_a$ (Figure 4.7). This residue appears to be important for stabilizing the RNA binding motif $KFRDI^{82-86}$ through its interaction with the backbone carbonyl of Arg^{84}. This interaction is seen for all the 3C picornaviridae proteases (HRV2-3C: Asn^{14} and Arg^{84}; Polio-3C: Asn^{14} and Arg^{84}; HAV-3C: Asn^{14} and Arg^{97}; FMD-3C: Asn^{15} and Arg^{97}). A second hydrogen bond is also observed between the side-chain amide of Asn^{14} and the backbone carbonyl of Gly^{29} for the HRV and Polio 3C proteases. Gly^{29} precedes the turn region His^{31} - Val^{34}, which has also been implicated in RNA recognition [32], and contains the conserved Asp^{32} residue that exhibits 'disallowed' backbone torsion angles.

Compared to the apo enzyme, the inhibited protein has a slightly increased RMSD in the region following the RNA binding motif (KFRDI), which has recently been implicated in heterodimerization with the 3D gene product, RNA polymerase [33]. Intermediate conformational exchange is also observed within this region. This is evident from the HNHA

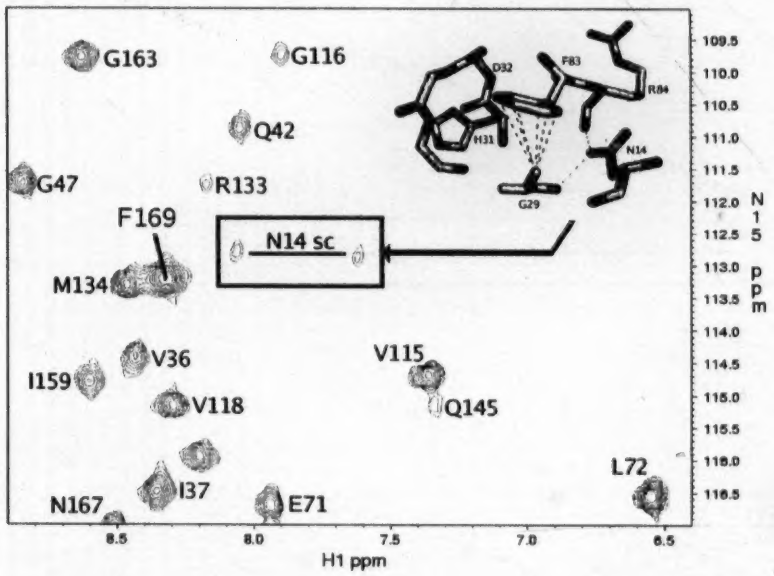

Figure 5.7: **Asn^{14} Side-chain Hydrogen Bonding in the Inhibited HRV14-3C Protease.** The deuterium exchange ^{15}N-HSQC spectrum collected ~2 hours after ^{1}H/^{2}H exchange of the inhibited HRV14-3C protease shows the inter-domain stabilization hydrogen bonds provided by the side-chain amide of Asn^{14}. The backbone carbonyl atoms from Phe^{83} and Gly^{29} act as hydrogen bond acceptors.

spectra shown in Figure 5.8. However, improved resolution in the ^{13}C-NOESY-HSQC spectrum was achieved for the apo enzyme compared to the inhibited enzyme because of the increased field strength (800 MHz) used to collected this experiment. This additionally improved both the restraint calibration and the structural calculation for this region. Slight differences for the backbone torsion angles in this region resulted and are believed to be more accurate in the apo enzyme (2IN2.pdb) relative to the bound state (2B0F.pdb).

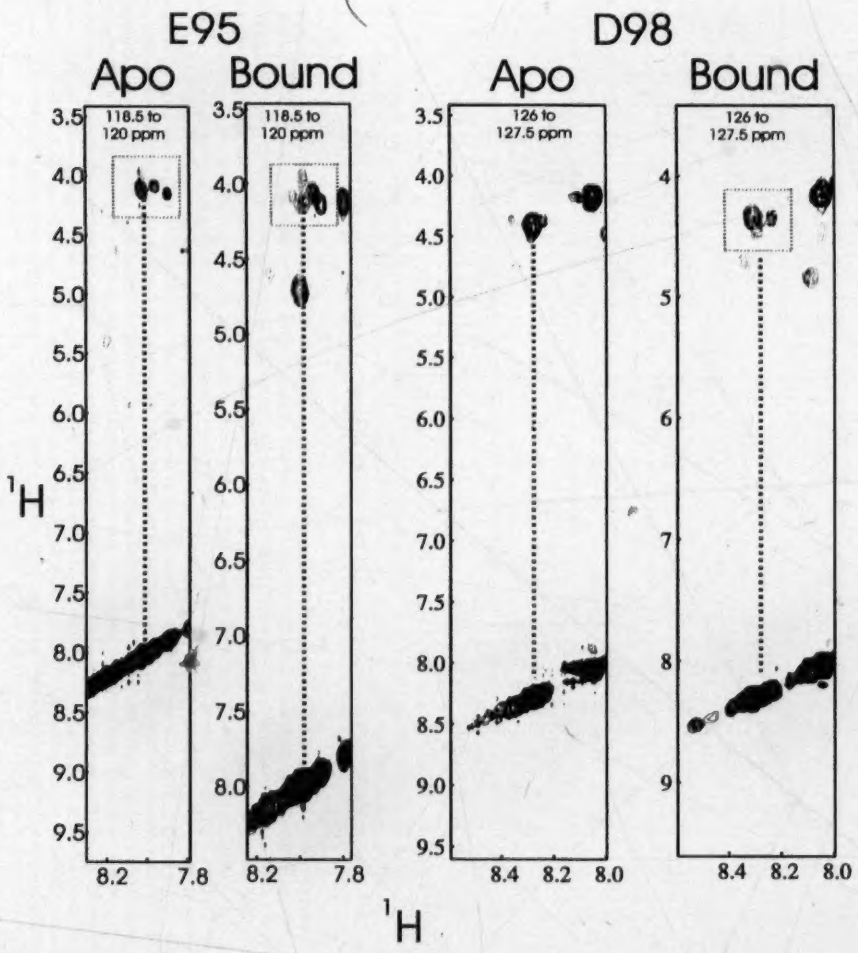

Figure 5.8: **HNHA Strip Plots for the Apo and Inhibited HRV14-3C Protease.** Strip plots for residues in the HRV14-3C protease implicated in binding the 3D gene product. Intermediate exchange exists for residue Glu^{95} in both states, however, persits for Asp^{98} in the bound state only. The peak broadening is boxed.

5.4.3 Deuterium Exchange and HRV14-3C Dynamics

The deuterium exchange data for the inhibited HRV14-3C protease shows that hydrogen exchange rates are similar and uniformly distributed between the two β-barrel domains and the RNA binding site (Figure 5.9).

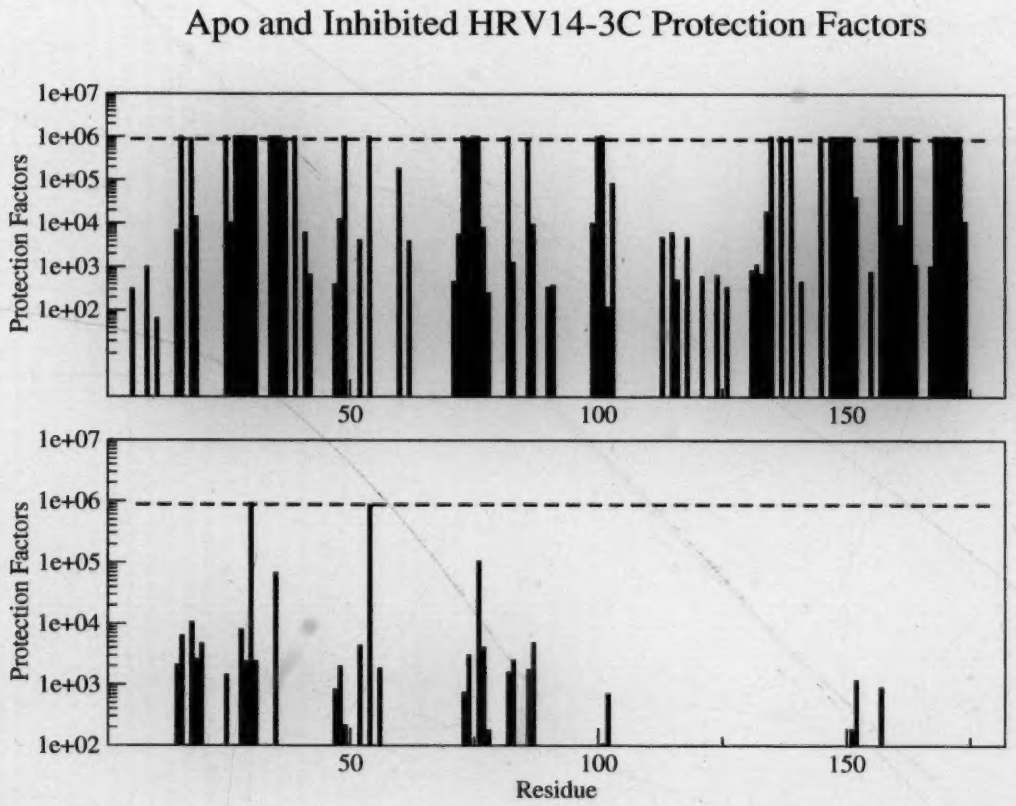

Figure 5.9: **Apo vs. Bound P_{factor} Data**. Slow exchange $^1H/^2H$ differences between the two states of the HRV14-3C protease are evident. The apo form exhibits a substantial global increase in flexibility. The C-terminal β-barrel domain, which interacts with proteolytic substrates, has no measurable P_{factor}s except for residues implicated with RNA binding [27]. The cut-off lines in each graph represent the maximum measurable protection factors that could be calculated during the 4 to 5 day experiment.

Fifty-five of the 87 residues involved in hydrogen bonds still had signals at the end of the deuterium exchange experiment, one week following $^1H/^2H$ exchange (Figure 5.10). These results show the overall structural uniformity and relative rigidity of the inhibited enzyme. Not surprisingly, amides that showed the largest protection factors are located within the interior of the protein and are involved in β -strand hydrogen bond interactions. Most of these slow-exchanging amides provided signals for months following the exchange process

and were of sufficient intensity to calculate T_2 values from the ^{15}N CPMG-HSQC spectra. Areas that demonstrate fast deuterium exchange rates are the loop/turn regions Lys^{61} - Asn^{69} and Ser^{128} - Thr^{131}, which interface the two domains; Arg^{79} - Glu^{81} and Glu^{92} - Asp^{98}, which flank the RNA binding site and His^{104} - Asn^{110}, which have been implicated in post-scissile bond substrate interactions. These regions correspond with areas providing higher than average normalized-atomic-displacement values (Figure 5.6).

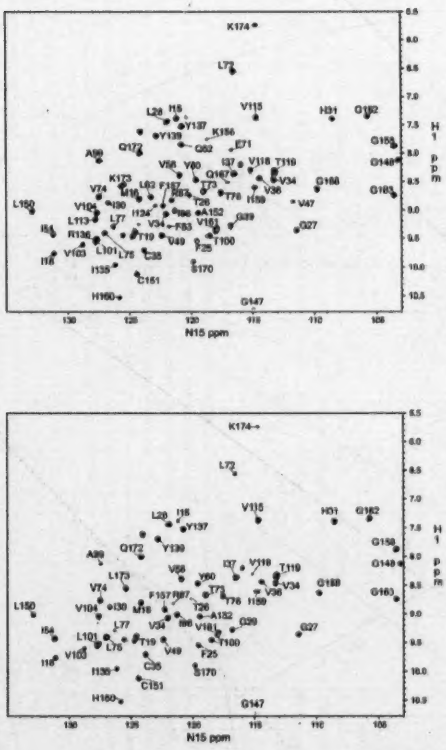

Figure 5.10: **Deuterium Exchange ^{15}N-HSQC Spectra of the Inhibited HRV14-3C Protease**. Two ^{15}N-HSQC spectra are shown, which were used to calculate deuterium exchange rates and P_{factor}s for the inhibited HRV14-3C protease. Fifty-five of the original 87 amide signals remained at the conclusion of the experiment.

Visual inspection of the D_2O exchange ^{15}N-HSQC spectra obtained for the apo (Figure 5.11) and the inhibited (Figure 5.10) forms of the HRV14-3C protease reveal some striking differences. About half the number of residues that presented amide signals in the inhibited form were visible following $^1H/^2H$ exchange of the apo enzyme. Comparison of the K_{ex} and P_{factor} data for the two states of HRV14-3C is provided in Table 5.4. The comprehensive calculations of this data are provided in *Appendix E*. The P_{factor} data for the two enzyme states are graphically represented in Figure 5.9. This figure illustrates the exchange dynamic differences between the two enzyme states. Except for the residues Cys^{151}, Ala^{152} and Lys^{155}, which are buried and are near residues implicated in RNA binding ($TGK^{154-156}$) in the homologous Polio-3C protease [27], the C-terminal β-barrel domain is devoid of any slow exchanging amides. Additionally, the measurable P_{factor}s for the apo enzyme are reduced by ~10 to 100-fold compared to the inhibited enzyme.

These data indicate that the C-terminal β-barrel domain undergoes faster amide exchange in the apo enzyme compared to its inhibited counterpart. Furthermore, it is possible that this domain may also exhibit increased flexible compared with its inhibited counterpart, which may be required to accommodate induced fitting of the various substrates.

Another area that demonstrated dynamic differences was the RNA binding region (His^{31} and KFRDI$^{82-86}$). These calculated K_{ex} data are displayed in Figure 5.12. The residues involved in this motif, namely Arg^{82} and Ile^{86}, present no change in peak volume throughout the $^{1}H/^{2}H$ exchange experiment for the inhibited form of the protease. This differs for the apo enzyme, in which conformational flexibility affords hydrogen exchange.

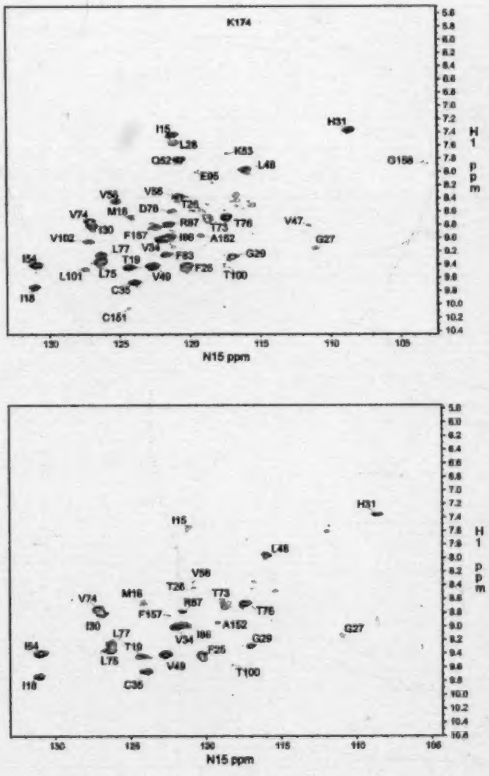

Figure 5.11: **Deuterium Exchange ^{15}N-HSQC Spectra of the Apo HRV14-3C.** Two ^{15}N-HSQC spectra used in the calculation of K_{ex} and P_{factor} data shown in Table 5.4 are shown. The number of residues presenting amide signals are $\sim \frac{1}{2}$ compared to the bound enzyme.

The backbone chemical shift analysis between the apo and and bound HRV14-3C enzymes (Figure 4.9) suggested no changes in backbone torsion angles for the RNA binding site upon inhibitor binding. These results were confirmed following the structure calculation of the apo enzyme that provided superposable backbone atoms (Table 5.3). This is interesting because bi-directional allosteric communication between the proteolytic and RNA binding sites was reported for the homologous HAV-3C protease by Peters *et al.* [34]. This group proposed two possible mechanisms for the communication. If their findings are reiterated with the rhinovirus 3C enzymes, the results reported here suggest that this communication may not facilitated

through a structural change of the RNA binding site. The P_{factor} calculations suggest a dynamic change, either solely or in conjunction with dimerization (homo or hetero), might be the medium for allosteric communication.

Although further structural work with the interacting RNA fragment will be needed to confirm this, the necessary chemical shift and structural data reported here compliment the corresponding RNA solution structure data (PDB codes 1RFR and 1TXS) and allows these studies to commence.

Table 5.4: **Apo and Inhibited HRV14-3C K_{ex} and P_{factor} Data**[†]

Residue	K_{ex} s^{-1}		P_{factor}	
	Inhibited	Apo	Inhibited	apo
F^{6}	9.15^{-3}	nd	1.44×10^{4}	nd
S^{9}	5.27^{-3}	nd	8.86×10^{4}	nd
L^{11}	2.75^{-2}	nd	1.91×10^{3}	nd
I^{15}	4.00^{-4}	1.31^{-3}	3.14×10^{5}	9.62×10^{3}
M^{16}	$<1.90^{-6}$	4.90^{-4}	$>1.00 \times 10^{6}$	3.31×10^{5}
I^{18}	$<1.90^{-6}$	1.60^{-4}	$>1.00 \times 10^{6}$	2.92×10^{5}
T^{19}	2.20^{-4}	1.18^{-3}	7.36×10^{5}	1.37×10^{5}
T^{20}	nd	7.50^{-4}	nd	2.84×10^{5}
F^{25}	$<1.90^{-6}$	1.94^{-3}	$>1.00 \times 10^{6}$	6.80×10^{4}
T^{26}	4.20^{-4}	nd	7.52×10^{5}	nd
G^{27}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
L^{28}	$<1.90^{-6}$	3.40^{-4}	$>1.00 \times 10^{6}$	3.71×10^{5}
G^{29}	$<1.90^{-6}$	1.96^{-3}	$>1.00 \times 10^{6}$	1.89×10^{5}
I^{30}	$<1.90^{-6}$	$<2.50^{-6}$	$>1.00 \times 10^{6}$	$>1.00 \times 10^{6}$
H^{31}	$<1.90^{-6}$	1.24^{-3}	$>1.00 \times 10^{6}$	1.22×10^{5}
V^{34}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
C^{35}	$<1.90^{-6}$	1.10^{-4}	$>1.00 \times 10^{6}$	8.87×10^{6}
V^{36}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
I^{37}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
T^{39}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
A^{41}	8.30^{-4}	nd	5.37×10^{5}	nd
Q^{42}	6.99^{-3}	nd	5.31×10^{4}	nd
V^{47}	4.15^{-3}	1.94^{-3}	1.03×10^{4}	2.20×10^{4}
L^{48}	1.60^{-4}	9.80^{-4}	3.85×10^{5}	6.29×10^{4}
V^{49}	$<1.90^{-6}$	7.20^{-3}	$>1.00 \times 10^{6}$	5.51×10^{3}
Q^{52}	1.36^{-3}	1.29^{-3}	4.03×10^{5}	4.26×10^{5}
I^{54}	$<1.90^{-6}$	$<2.50^{-6}$	$>1.00 \times 10^{6}$	$>1.00 \times 10^{6}$
V^{56}	nd	1.38^{-3}	nd	7.75×10^{4}
Y^{60}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
L^{62}	6.50^{-4}	nd	1.72×10^{5}	nd
E^{71}	4.18^{-3}	nd	1.47×10^{4}	nd
L^{72}	3.40^{-4}	nd	1.77×10^{5}	nd
T^{73}	$<1.90^{-6}$	4.26^{-3}	$>1.00 \times 10^{6}$	3.99×10^{5}
V^{74}	$<1.90^{-6}$	5.70^{-4}	$>1.00 \times 10^{6}$	8.87×10^{4}
L^{75}	$<1.90^{-6}$	nd	$>1.00 \times 10^{6}$	nd
T^{76}	$<1.90^{-6}$	3.00^{-5}	$>1.00 \times 10^{6}$	5.67×10^{6}
L^{77}	2.50^{-4}	4.90^{-4}	2.64×10^{5}	1.35×10^{5}
D^{78}	9.56^{-3}	1.37^{-2}	1.04×10^{4}	7.27×10^{3}
K^{82}	$<1.90^{-6}$	2.17^{-3}	$>1.00 \times 10^{6}$	9.63×10^{4}
F^{83}	2.96^{-3}	1.50^{-3}	8.28×10^{4}	1.63×10^{5}
I^{86}	$<1.90^{-6}$	8.90^{-4}	$>1.00 \times 10^{6}$	4.46×10^{4}
R^{87}	3.90^{-4}	7.60^{-4}	5.86×10^{5}	9.09×10^{4}
I^{90}	6.00^{-3}	nd	1.15×10^{4}	nd

Continued on Next Page...

Table 5.4: Apo and Inhibited HRV14-3C K_{ex} and P_{factor} Data – Continued

Residue	K_{ex} s^{-1}		P_{factor}	
	Inhibited	Apo	Inhibited	Apo
S^{91}	1.34^{-2}	nd	3.32×10^4	nd
A^{99}	3.50^{-4}	nd	6.10×10^5	nd
T^{100}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
L^{101}	$<1.90^{-6}$	1.87^{-2}	$>1.00 \times 10^6$	3.53×10^3
V^{102}	1.30^{-2}	2.20^{-3}	3.07×10^3	1.80×10^4
V^{103}	2.00^{-5}	nd	2.33×10^6	nd
L^{113}	3.60^{-4}	nd	1.39×10^5	nd
V^{115}	2.70^{-4}	nd	1.69×10^5	nd
G^{116}	9.82^{-3}	nd	4.44×10^4	nd
V^{118}	3.20^{-4}	nd	1.16×10^5	nd
A^{121}	8.09^{-3}	nd	5.14×10^4	nd
I^{124}	2.33^{-3}	nd	1.59×10^4	nd
L^{126}	9.55^{-3}	nd	1.86×10^4	nd
T^{131}	3.69^{-3}	nd	4.29×10^4	nd
N^{132}	5.91^{-3}	nd	1.31×10^5	nd
R^{133}	9.85^{-3}	nd	8.24×10^4	nd
M^{134}	2.90^{-4}	nd	1.81×10^6	nd
I^{135}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
Y^{137}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
Y^{139}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
T^{141}	8.90^{-3}	nd	3.09×10^4	nd
Q^{145}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
G^{147}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
G^{148}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
V^{149}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
L^{150}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
C^{151}	$<1.90^{-6}$	3.75^{-2}	$>1.00 \times 10^6$	2.21×10^4
A^{152}	2.00^{-4}	6.90^{-3}	5.73×10^6	1.67×10^5
K^{155}	6.47^{-3}	nd	6.74×10^4	nd
F^{157}	$<1.90^{-6}$	2.83^{-3}	$>1.00 \times 10^6$	3.85×10^4
G^{158}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
I^{159}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
H^{160}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
V^{161}	2.50^{-4}	nd	3.56×10^5	nd
G^{162}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
G^{163}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
N^{164}	8.10^{-3}	nd	1.82×10^5	nd
Q^{167}	5.71^{-3}	nd	1.08×10^5	nd
G^{168}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
F^{169}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
S^{170}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
A^{171}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
Q^{172}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
L^{173}	$<1.90^{-6}$	nd	$>1.00 \times 10^6$	nd
K^{174}	3.00^{-4}	nd	6.06×10^5	nd

†nd = Not Determined (fast)

5.4.4 Oligomerization Analysis

During the analysis of the backbone NMR data for the apo HRV14-3C protease, the possibility that the enzyme was oligomerizing was investigated. These concerns originated

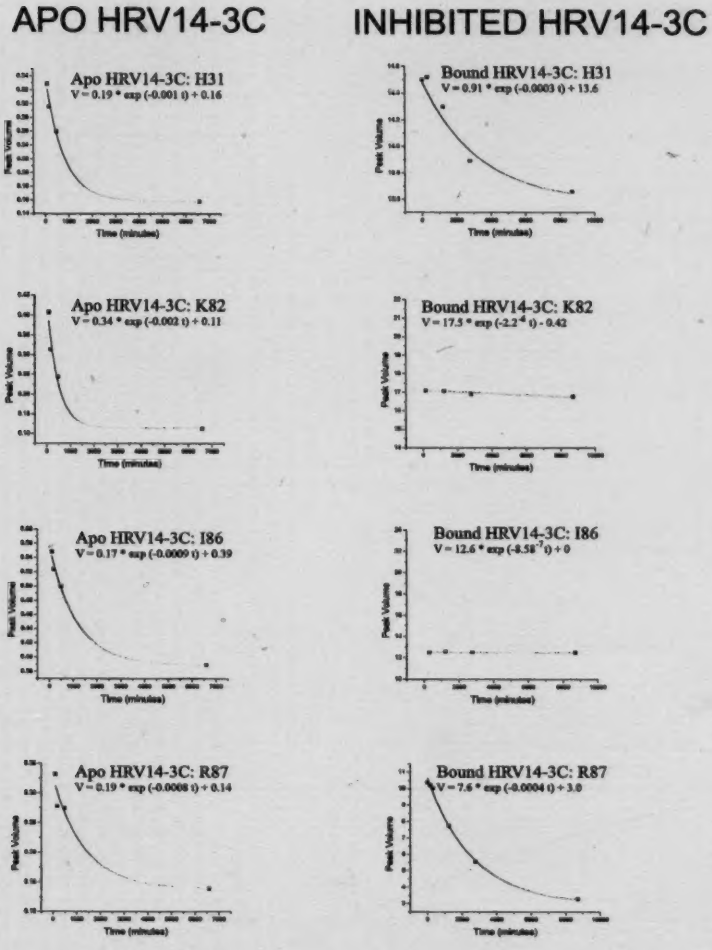
APO HRV14-3C
INHIBITED HRV14-3C
Apo HRV14-3C: H31
V = 0.19 * exp (-0.001 t) + 0.16
Bound HRV14-3C: H31
V = 0.91 * exp (-0.0003 t) + 13.6
Apo HRV14-3C: K82
V = 0.34 * exp (-0.002 t) + 0.11
Bound HRV14-3C: K82
V = 17.5 * exp (-2.2^-4 t) - 0.42
Apo HRV14-3C: I86
V = 0.17 * exp (-0.0009 t) + 0.39
Bound HRV14-3C: I86
V = 12.6 * exp (-8.58^-7 t) + 0
Apo HRV14-3C: R87
V = 0.19 * exp (-0.0008 t) + 0.14
Bound HRV14-3C: R87
V = 7.6 * exp (-0.0004 t) + 3.0
Peak Volume
Time (minutes)

Figure 5.12: **Selected K_{ex} Data for the KFRDI Region of the HRV14-3C Protease.** Volume integration values were obtained from NMRView [9] and fitted to the exponential curves with ORIGIN v7.5. The selected graphs are for the RNA binding region KFRDI$^{82-86}$. Data for the apo enzyme is presented on the left. Data for the inhibited enzyme is shown on the right. Despite the near surface exposure of residues Lys^{82} and Ile^{86}, their amides present no measurable exchange in the bound state compared with the apo enzyme.

from preliminary ultracentrifugation data that was collected on the apo enzyme during the initial NMR sample condition screening, which fit a weak monomer-dimer or weak monomer-tetramer model with a predicted mass of 24 kDa. These results were confirmed *via* NMR. Fitting the integrated ^{15}N-CPMG-HSQC envelope data to equation 5.3 yielded a T_2 of 62.5 ms corresponding to a correlation time (τ_c) of 14.1 ns in D_2O or ~10.57 ns in H_2O as determined from equation 5.1 for the inhibited enzyme. The estimated τ_c (equation 5.4) for a monomeric form of the labelled HRV14-3C protease is 9.25 ns in H_2O and 11.56 ns in D_2O. These results are consistent with a weak dimeric model where ~86% of the enzyme exists as a monomer as determined with equation 5.6. The results for the apo enzyme in H_2O provided a T_2 of 75 ms and a τ_c of 12.1 ns. These results correspond with a weak dimeric model with ~70% of the apo enzyme existing as a monomer. Because comparison of this data involves samples in different solvents (H_2O and D_2O) and because an expected error of ~7% exists when comparing data obtained between the two solvent, the minor difference between the two samples (16%) is probably insignificant. These results were expected considering every attempt was undertaken to obtain and maintain a monomeric form of the protease to improve the quality of NMR data. Interstingly, two possible homodimerization interfaces were reported for the X-ray structure of the homologous Polio-3C protease [29, 35]. One interface involved residues Ile^{56} - Glu^{63} and the other involved residues Tyr^{109} Pro^{115}. These regions correspond to the regions Val^{56} - Asp^{64} and Thr^{109} - Val^{118} in the HRV14-3C protease. Broad amide NMR peaks were observed in the HRV14-3C protease for Val^{56}, which resides within one of the suggested homodimerization interfaces (Figure 4.7: β-strand $\mathbf{I}_e$). However, inspection of other residues within this β-strand did not reveal similar results and identification of any dimerization site could not be made with the NOE data collected on either the apo or inhibited HRV14-3C enzyme. Whether these exposed β-strand backbone atoms are involved in non-specific homodimerization or are specific functional heterodimerization interfaces necessary for translation and/or transcription regulation [33, 36] is still undetermined.

5.5 Conclusion

The preliminary concerns regarding sample stability were resolved following completion of this phase of the project. However, despite the additional purification step with the hydroxyapatite column and the addition of 0.1 mM NaN_3, the apo enzyme eventually precipitated out of solution (after ~6 months). Interestingly, the broad NMR peaks observed in the ^{15}N-HSQC spectrum (Figure 5.1) of the apo HRV14-3C protease were originally thought to arise from degradation. However, these observations were also seen in spectra collected on an apo HRV14-3C protease sample with significantly improved stability and therefore, postulated to result from μs chemical exchange processes. Unlike the inhibited enzyme, which presented nearly all expected NMR data, the apo enzyme has a number of regions with broad (Figure 5.3) or un-assignable signals that are hypothesized to be involved in conformational exchange processes. These areas are presumed to undergo μs timescale chemical changes. Interestingly, the majority of these residues fall within the proteolytic site, the stemloop-D RNA and RNA polymerase binding sites and in regions with yet undetermined biological functions. The apo enzyme also presented increased RMSDs in the C-terminal β-barrel domain compared to the inhibited enzyme. Although the few number of data points limited the precision of the calculated exchange rates, a clear distinction between the two domains was evident. Specifically, the C-terminal β-barrel domain, which is involved in proteolytic substrate recognition, presented faster ^{1}H/^{2}H exchange rates (Figure 5.4) in the apo enzyme compared with the inhibited enzyme. The backbone of the RNA binding region remains conformationally unchanged upon inactivation. Dynamically, however, it presents slower exchanging amides. These new insights compliment recent findings regarding allosteric communication between the RNA binding and proteolytic regions of the 3C protease, which was observed for the homologous enzyme, HAV-3C [34]. The findings presented here suggest that if these processes occur in the HRV14-3C enzyme, they are dynamically rather than structurally mediated.

Bibliography

[1] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N Chemical Shift Referencing in Biomolecular NMR. *J Biomol NMR*, 6:135–140, 1995.

[2] L. E. Kay, P. Keifer, and T. Saarinen. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J Am Chem Soc*, 114:10663–10665, 1992.

[3] M. Ikura, L. E. Kay, and A. Bax. A Novel Approach for Sequential Assignment of ^{1}H, ^{13}C, and ^{15}N Spectra of Larger Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin. *Biochemistry*, 29:4659–4667, 1990.

[4] M. Wittekind and L. Mueller. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha and Beta-Carbon Resonances in Proteins. *J Magn Reson B*, 101:201–205, 1993.

[5] S. Grzesiek and A. Bax. Correlating Backbone Amide and Side Chain Resonances in Larger Proteins by Multiple Relayed Triple Resonance NMR. *J Am Chem Soc*, 114:6291–6293, 1992.

[6] D. R. Muhandiram and L. E. Kay. Gradient-Enhanced Triple-Resonance 3-Dimensional NMR Experiments with Improved Sensitivity. *J Magn Reson B*, 103:203–216, 1994.

[7] G. W. Vuister and A. Bax. Quantitative J Correlation: A New Approach for Measuring Homonuclear Three-Bond J(HNHα) Coupling Constants in ^{15}N-Enriched Proteins. *J Am Chem Soc*, 115:7772–7777, 1993.

[8] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: a Multidimensional Spectral Processing System Based on UNIX Pipes. *J Biomol NMR*, 6:277–293, 1995.

[9] B. A. Johnson and R. B. Blevins. NMRView: A Computer Program for the Visualization and Analysis of NMR Data. *J Biomol NMR*, 4:603–614, 1994.

[10] L. Spyracopoulos, S. M. Gagne, and B. D. Sykes. *Proceedings for the International School of Structural Biology and Magnetic Resonance. 4th Course on Dynamics, Structure and Function of Biological Macromolecules.* Plenum Press, New York, NY, USA, 1999.

[11] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton. *Protein NMR Spectroscopy: Priciples and Practice.* Academic Press, New York, NY, USA, 1996.

[12] Y. Bai, J. S. Milne, L. Mayne, and S. W. Englander. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins*, 17:75–86, 1993.

[13] G. P. Connelly, Y. Bai, M. Jeng, and S.W. Englander. Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins*, 17:87–92, 1993.

[14] P. Güntert, C. Mumenthaler, and K. Wüthrich. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol*, 273:283–298, 1997.

[15] T. Herrmann, P. Güntert, and K. Wüthrich. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol*, 319:209–227, 2002.

[16] C. Mumenthaler, P. Güntert, W. Braun, and K. Wüthrich. Automated Combined Assignment of NOESY Spectra and Three-dimensional Protein Structure Determination. *J Biomol NMR*, 10:351–362, 1997.

[17] O. Zhang, L. E. Kay, J. P. Olivier, and J. D. Forman-Kay. Backbone ^{1}H and ^{15}N Resonance Assignments of the N-terminal SH3 Domain of Drk in Folded and Unfolded States Using Enhanced-Sensitivity Pulsed Field Gradient NMR Techniques. *J Biomol NMR*, 4:845–858, 1994.

[18] G. Cornilescu, F. Delaglio, and A. Bax. Protein Backbone Angle Restraints from Searching a Database for Chemical Shift and Sequence hHomology. *J Biomol NMR*, 13:289–302, 1999.

[19] D. A. Matthews, W. W. Smith, R. A. Ferre, B. Condon, G. Budahazi, W. Sisson, J. E. Villafranca, C. A. Janson, H. E. McElroy, C. L. Gribskov, and Worland S. Structure of Human Rhinovirus 3C Protease Reveals a Trypsin-like Polypeptide Fold, RNA-Binding Site, and Means for Cleaving Precursor Polyprotein. *Cell*, 77:761–771, 1994.

[20] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, Read R. J., L. M. Rice, T. Simonson, and G. L. Warren. Crystallography & NMR system: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr D Biol Crystallogr*, 54:905–921, 1998.

[21] A. J. Nederveen, J. F. Doreleijers, W. Vranken, Z. Miller, C. A. E. M. Spronk, S. B. Nabuurs, P. Guentert, M. Livny, J. L. Markley, M. Nilges, Ulrich E. L., R. Kaptein, and A. M. Bonvin. RECOORD: A Recalculated coordinate Database of 500+ Proteins from the PDB Using Restraints from the BioMagResBank. *Proteins*, 59:662–672, 2005.

[22] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thorton. PROCHECK: A Program to check the Stereochemical Quality of Protein Structures. *J Appl Cryst*, 26:283–290, 1993.

[23] R. W. W. Hooft, G. Vriend, C. Sander, and E. E. Abola. Errors in Protein Structures. *Nature*, 381:272, 1996.

[24] R. Koradi, M. Billeter, and K. Wüthrich. MOLMOL: A Program for Display and analysis of Macromolecular Structures. *J Mol Graphics*, 14:51–55, 1996.

[25] L. Willard, A. Ranjun, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes, and Wishart D. S. VADAR: A Web Server for Quantitative Evaluation of Protein structure Quality. *Nucleic Acids Res*, 31:3316–3319, 2003.

[26] J.F. Doreleijers, M.L. Raves, J.A.C. Rullmann, and R. Kaptein. Completeness of NOEs in Protein Structure: A Statistical Analysis of NMR. *J Biomol NMR*, 14:123–132, 1999.

[27] S-R. Shih, C. Chiang, T-C. Chen, C-N. Wu, J. T-A. Hsu, J-C. Lee, M-J. Hwang, M-L. Li, G-W. Chen, and M-S. Ho. Mutations at KFRDI and VGK Domains of Enterovirus 7C Protease Affect Its RNA Binding and Proteolytic Activities. *J. Biomed 'Science*, 11:239–248, 2004.

[28] Y Yang, R. Rijnbrand, S. Watowich, and S. W. Lemon. Genetic Evidence for an Interaction Between a Picornaviral *cis*-Acting RNA Replication Element and 3CD Protein. *J Biol Chem*, 279:12659–12667, 2004.

[29] S. C. Mosimann, M. M. Cherney, S. Sia, S. Plotch, and M. N. G. James. Refined X-ray Crystallographic Structure of the Poliovirus 3C Gene Product. *J Mol Biol*, 273:1032–1047, 1997.

[30] M. Billeter, A. D. Kline, W. Braun, R. Huber, and K. Wüthrich. Comparison of the High-Resolution Structures of the Alpha-Amylase Inhibitor Tendamistat Determined by Nuclear Magnetic Resonance in Solution and by X-ray Diffraction in Single Crystals. *J Mol Biol*, 206:677–687, 1989.

[31] M. Billeter. Comparison of Protein Structures Determined by NMR in Solution and by X-ray Diffraction in Single Crystals. *Q Rev Biophys*, 25:325–377, 1992.

[32] R. Andino, G. E. Rieckhof, P. L. Achacoso, and D. Baltimore. Poliovirus RNA Synthesis Utilizes an RNP Complex Formed Around the 5′-end of Viral RNA. *Embo J*, 12:3587–3598, 1993.

[33] D. W. Yang, Y. Zheng, D. Liu, and D. F. Wyss. Sequence-Specific Assignments of Methyl Groups in High-Molecular Weight Proteins. *J Am Chem Soc*, 126:3710–3711, 2004.

[34] H. Peters, Y. Y. Kusov, S. Meyer, A. J. Benie, E. Bauml, M. Wolff, C. Rademacher, T. Peters, and V. Gauss-Muller. Hepatitis A Virus Proteinase 3C Binding to Viral RNA: Correlation with Substrate Binding and Enzyme Dimerization. *J Biochem*, 385:363–370, 2005.

[35] E. M. Bergmann, S. C. Mosimann, M. M. Chernaia, B. A. Malcolm, and M. N. G. James. The Refined Crystal Structure of the 3C Gene Product from Hepatitis A Virus: Specific Proteinase Activity and RNA Recognition. *J Virol*, 71:2436–2448, 1997.

[36] A. V. Gamarnik and R. Andino. Switch from Translation to RNA Replication in a Positive-Stranded RNA Virus. *Genes and Dev*, 12:2293–2304, 1998.

Chapter 6

Concluding Remarks and Future Directions

6.1 Introduction

The intent of this study was to investigate the proteolytic pharmacophore of the rhinovirus 3C protease (serotype 14, subgenus B) using NMR methodology. While conducting the initial NMR experiments on the apo enzyme however, solution state instability problems persisted that made NMR spectral analysis difficult. Therefore, it was determined that in order to complete this objective, further stabilization of the enzyme would be required. This led to the investigation of an inhibited state in which the enzyme was inactivated by covalently binding a 2C/3A peptide cleavage sequence analogue. This change in direction greatly improved the enzyme's solution stability and allowed the exploration of the pharmacophore with a novel inhibitor that deepened our understanding of 3D protease substrate diversity. These insights were obtained by comparing the solution structure of HRV14-3C with the X-ray structure of HRV2-3C (subgenus A; 51% sequence identity to the HRV14-3C protease). In addition to this comparison, charaterization of previously unknown upstream (P_5 and P_6) substrate interactions were afforded from the increased length of our inhibitor. Not only has our understanding of the conserved components of the HRV-protease substrate binding mechanism been improved, but new insights into the deactivation mechanism of the enzyme *via* deamidation of Asn^{164} has also been gained.

This project also succeeded in producing the solution structure for the apo HRV14-3C enzyme along with a nearly complete chemical shift assignment. Investigation into dynamic

differences between the two enzyme states yielded some interesting results. Specifically, the C-terminal β-barrel domain presents fast exchanging amides in comparison to the inhibitor bound state. Upon binding of the inhibitor, the entire enzyme becomes 'rigid'. This inactivation also manifested dramatic improvements to solution stability of the protease.

There are two directions for future consideration that this project might take. One centers around answering additional biological questions. The other continues with the development of inhibitors and potential lead pharmaceutical candidates.

6.2 Biological Questions

6.2.1 Heterodimerization Interface Identification

In completing this project, two structures and two sets of chemical shift data for two different states of the HRV14-3C protease were produced. These data now provided the necessary foundation to explore other biological processes of the enzyme through chemical shift perturbation analysis. By using NMR as a tool for structure elucidation, previously identified homo-dimerization sites in a homologous 3C protease [1] could not be confirmed in the solution structure of the HRV14-3C protease under the sample conditions outlined in Chapter 2. However, RMSD changes and dynamic differences (Table 5.4) identified three key regions of the protein. Two of the three regions have already been confirmed to be involved with important biological functions. One region being the proteolytic site (studied here), the other region being the RNA binding site [2] and the last region being an area implicated with viral RNA polymerase heterodimerization [3]. What is noteworthy from this work, is that additional regions of the protease also display changes in their dynamics (*ie.* Val^{56}, Lys^{61}, Leu^{62}, Asn^{67} and Asn^{80}). Although spatially close, these residues have not been implicated with any specific hetero-dimerization interface. The important questions to be asked here, are what other proteins does the 3C enzyme associate with? The 3C protease has been shown to be regulated by the 2C gene product. Also, the 3C protease has been implicated with altering host-cell RNA transcription machinery, as have the 3AB gene products. Does the 3C protease interface with any of these proteins and if so, where? These are questions that might be answered from simple ^{15}N-HSQC titration experiments relatively quickly now that the chemical shift assignments are in hand.

Furthermore, two important biological processes that HRV14-3C is involved with (proteolytic activity and RNA binding/RNA polymerase heterodimerization) have been allosterically linked for the homologous picornaviral 3C protease, HAV-3C [4]. It is still unknown if these effects manifest in the rhinovirus proteases, which leaves a door open for these studies to be explored kinetically. Following this, structural studies could also be undertaken. Interstingly, no structural change in the RNA binding site occured upon inhibition. Determining whether structural changes occur in the proteolytic site following RNA binding would allow us to determine if allosteric communication exists.

6.3 Inhibitor Design

6.3.1 Proteolytic Inhibitors

It was hoped that the insights into the proteolytic recognition gained with this study could help direct future inhibitor design. Two points were concluded from this study. One being that lead pharmaceutical candidates should be truncated to the P_4 position. This conclusion arose because inhibitors with only the P_1 to P_4 substrate positins have already been shown to impart substrate recognition and furthermore, the Asn^{164} deamidation event has been shown to significantly impair this recognition [5]. By limiting the size of inhibitors to the P_3 or P_4 upstream position, no interaction with the asparagine would occur, making the deamidation event a mute point. The second issue focuses around the P' substrate interactions. No structural data for these interactions exists for any of the HRV-3C proteases. Inclusion of residues that bind to this substrate pocket have presented significantly increased cleavage rates. It is hypothesized that inhibitors that incorporate post-scissile bond moieties into their design, might compensate for the loss of substrate recognition afforded by the P_5 to P_8 residues. The studies described here have identified a region of the 3C protease that might well be involved with these post-scissle bond interactions. Although no structural data involving these post-scissile bond interactions yet exists, the proximity of these residues to the catalytic triad and their strong correlation with the apo enzyme in regards to flexibility and dynamic similarity suggest their possible involvement. Future projects might focus around characterizing these interactions. This information could help develop inhibitors that bind the P_3 to P'_3 substrate positions.

6.3.2 Inhibition with Zinc

Other sets of experiments might focus around studying the possible inactivation of the enzyme with zinc. Since its implication in expediting the recovery of individuals infected with rhinoviruses.[6], zinc has been the focus of a number of double-blind studies to test its effectiveness in alleviating cold symptoms [7, 8, 9]. These studies have shown a statistically significant difference between the control and zinc treated groups. These results are intersting given the fact the zinc binds cysteine residues and the 3C gene product is a cysteine protease. Does zinc interact within the proteolytic active site or the other two cysteine residues (Cys^{35} and Cys^{151}), which are spatially close to the RNA and 3D heterodimeration interfaces? This question might also be answered from a simple ^{15}N-HSQC titration experiment.

6.3.3 Other Methonds of Inhibition

The philosophy "*there are more than one way to skin a cat* " might hold true for inhibiting the 3C protease. In particular, sequential and structural conservation is observed within the 3C protease proteolytic and RNA binding sites (Figure 6.1). Although their catalytic triad residues are conserved, residues in proximity to the triad that are involved with binding the multitude of proteolytic substrates are divergent within the picornaviruses. In fact, the 110+ serotypes within the HRV genus alone exhibit considerable variability in their substrate and inhibitor recognition [10]. This variability had made the development of universal proteolytic inhibitors based on active-site inhibition somewhat challenging. However, members within the picornaviridae 3C protease family also share a sequentially and structurally conserved RNA binding region (Table 5.3).

The binding determinant of this RNA fragment (stem loop D of IRES 1 clover-leaf RNA), which binds the picornaviral 3C protease has been identified [12] and its solution structure solved by NMR [13, 14]. Furthermore, our study shows that the RNA binding region of the HRV14-3C protease remains rigid upon proteolytic inactivation. These results mirror the results seen with the apo and bound structures reported for the homologous hepatitis A (HAV-3C) protease (Table 5.3). The high level of conservation seen for the picornaviral 3C protease RNA binding site suggests that the development of competitive

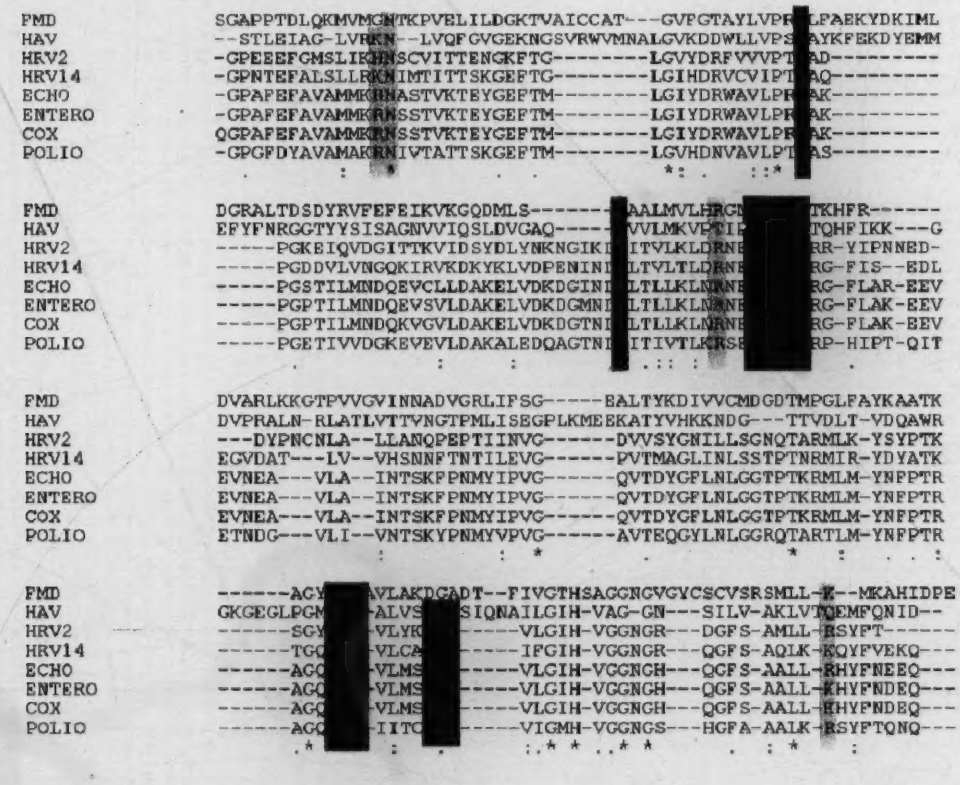

Figure 6.1: **Picornaviridae Protease Sequential Alignment**. Multiple sequence alignment for various picornavirus 3C proteases was performed with CLUSTLW [11]. Sequence conservation of the RNA binding site is highlighted red. Residues with suspected RNA interaction are colored brown and the residues implicated in 3D polymerization are colored yellow. The proteolytic triad and oxyanion-hole forming residues are highlighted green.

RNA binding antagonists might also be possible. In addition, these inhibitors may also afford universal applications in the treatment of other picornaviridae diseases, which include more pathogenic organisms like Polio, Hepatitis-A, Meningitis, Myocarditis and Foot-and-mouth disease. Certainly, more investigation into the structural similarity of residues in proximity to the RNA binding site will need to be done to confirm this and no structural data yet exists for the protein-RNA binding. However, now that structural and chemical shift data is in hand for both of these macromolecules, this study can proceed.

It might be possible that inhibition of other protein-protein interactions may also impact in the viral life cycle significantly. For example, would destroying the 3CD heterodimer-

ization event halt the nuclear localization of the 3C protease or affect RNA transcription? What other interface sites exist (*ie.* 2C gene product)? First and foremost, before any of these studies can begin, these possible interface sites need to be identified and characterized structurally.

6.4 Conclusion

Since beginning this study, many new insights into the picornaviral life cycle have emerged. Our understanding of how the virus regulates its activities, affects host cell activities and eventually replicates have been improved. This thesis presents new data about a very specific component required by the picornaviridae, which can now be added to this growing body of knowledge. Ultimately, it is hoped that the insights provided here can be used to further our scientific understanding of the virus' life cycle and be used to help direct other researchers in their quest to interupt the virus' capacity to infect and replicate. For this outcome is what it truly sought and what is required to reduce human morbidity and improve human quality of life.

Bibliography

[1] S. C. Mosimann, M. M. Cherney, S. Sia, S. Plotch, and M. N. G. James. Refined X-ray Crystallographic Structure of the Poliovirus 3C Gene Product. *J Mol Biol*, 273:1032–1047, 1997.

[2] S-R. Shih, C. Chiang, T-C. Chen, C-N. Wu, J. T-A. Hsu, J-C. Lee, M-J. Hwang, M-L. Li, G-W. Chen, and M-S. Ho. Mutations at KFRDI and VGK Domains of Enterovirus 7C Protease Affect Its RNA Binding and Proteolytic Activities. *J. Biomed 'Science*, 11:239–248, 2004.

[3] D. W. Yang, Y. Zheng, D. Liu, and D. F. Wyss. Sequence-Specific Assignments of Methyl Groups in High-Molecular Weight Proteins. *J Am Chem Soc*, 126:3710–3711, 2004.

[4] H. Peters, Y. Y. Kusov, S. Meyer, A. J. Benie, E. Bauml, M. Wolff, C. Rademacher, T. Peters, and V. Gauss-Muller. Hepatitis A Virus Proteinase 3C Binding to Viral RNA: Correlation with Substrate Binding and Enzyme Dimerization. *J Biochem*, 385:363–370, 2005.

[5] G. A. Cox, R. B. Johnson, J. A. Cook, M. Wakulchik, M. G. Johnson, E. C. Villarreal, and Q. M. Wang. Identification and Characterization of Human Rhinovirus-14 3C Protease Deamidation Isoform. *J Biol Chem*, 274:13211–13216, 1999.

[6] B. D. Korant, J. C. Kauer, and B. E. Butterworth. Zinc Ions Inhibit Replication of Rhinoviruses. *Nature*, 248:588–590, 1974.

[7] D R Davis G A Eby and W W Halcomb. Reduction in Duration of Common Colds by Zinc Gluconate Lozenges in a Double-blind Study. *Antimicrob Agents and Chemo*, 25:20–24, 1984.

[8] B. M. Farr, E. M. Conner, R. F. Betts, J. Oleske, A. Minnefor, and Gwaltney Jr. J. M. Two Randomized Controlled Trials of Zinc Gluconate Lozenge Therapy of Experimentally Induced Rhinovirus Colds. *Antimicrob Agents and Chemo*, 31:1183–1187, 1987.

[9] S. B. Mossad, M. L. Macknin, S. V. Mendendorp, and P. Mason. Zinc Gluconate Lozenges for Treating the Common Cold: A Randomized, Double-Blind, Placebo-Controlled Study. *Annals of Intern Med*, 125:81–88, 1996.

[10] T. O. Johnson, Y. Hua, H. T. Luu, E. L. Brown, F. Chan, S. S. Chu, P. S. Dragovich, B. W. Eastman, R. A. Ferre, S. A. Fuhrman, T. F. Hendrickson, F. C. Maldonado,

D. A. Matthews, J. W. Meador III, A. K. Patick, S. H. Reich, D. J. Skalitzky, S. T. Worland, M. Yang, and L. S. Zalman. Structure-Based Design of a Parallel Synthetic Array Directed Toward the Discovery of Irreversible Inhibitors of Human Rhinovirus 3C Protease. *J Med Chem*, 45:2016–2023, 2002.

[11] D. Higgins, J. Thompson, T. Gibson, J. D. Thompson, D. G. Higgins, and Gibson T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res*, 22:4673–4680, 1994.

[12] P. A. Walker, L. E-C. Leong, and A. G. Porter. Sequence and Structural Determinants of the Interaction between the 5'-Noncoding Region of Picornavirus RNA and Rhinovirus Protease 3C. *J Biol Chem*, 270:14510–14516, 1995.

[13] O. Ohlenschlager, J. Wohnert, E. Bucci, S. Seitz, S. Hafner, R. Ramachandran, R. Zell, and M. Gorlach. The structure of the stemloop D subdomain of coxsackievirus B3 cloverleaf RNA and its interaction with the proteinase 3C. *Structure*, 12:237–248, 2004.

[14] Z. Du, J. Yu, N. B. Ulyanov, and T. L. James. Solution Structure of a Consensus Stem-Loop D RNA Domain thay Plays Important Roles in Regulation Translation and Replication in Enteroviruses and Rhinoviruses. *Biochemistry*, 43:11959–11972, 2004.

Appendix A

Molecular Biology Recipes

Chromatography Buffer

Ingredient	Quantity
TRIS	6.05 G
EDTA	2 ml
DTT	771 mg
dd H_2O at 4 °C	qs 1 L

Dissolve ingredients in dd H_2O and adjust the pH to 8.8 - 9.0 at 4 °C. with 10 M HCl (~45 drops at 4 °C). Trickled argon gas through the buffer for 5 minutes prior to adding DTT.

Denaturing Buffer

Ingredient	Concentration
Urea	7 M
TRIS-HCl pH 8.0	30 mM
Cysteine	20 mM
EDTA	1 mM
DTT	25 mM

Add the precipitated HRV14-3C protease up to a concentration a 1 mg/ml.

Refolding Buffer

Ingredient	Concentration
KH_2PO_4 pH 6.5	20 mM
EDTA	0.5 mM
DTT	5 mM

Equilibrate to 4 °C and degassed with Argon prior to addition of DTT.

Electrophoresis Running Buffer 10x

Ingredient	Quantity
Glycine	144 G
TRIS HCl	30 G
SDS	10 G
dd H_2O	qs 1 L

Dilute 1:10 prior to use.

Electrophoresis Staining Solution

Ingredient	Quantity
Glacial acetic acid	100 ml
Ethanol 95%	500 ml
Brilliant Blue R250	1 G
dd H_2O	400 ml

Electrophoresis De-Staining Solution

Ingredient	Quantity
Glacial acetic acid	400 ml
Ethanol 95%	1 L
dd H_2O	2.6 L

LB Plates

Ingredient	Quantity
Bactotryptan (Tryptone Peptone)	10 G
Yeast extract	5 G
NaCl	10 G
Bacto-Agar	15 G
dd H_2O	960 ml
Ampicillin	100 mg
Chloramphenicol	50 mg

Dissolve all ingredient except antibiotics in dd H_2O, and adjusted the pH to 7.0 with 1 M NaOH. Autoclave the media for 30 minutes at 15 psi and 121 °C. Once cooled below 50 °C, add the antibiotics and transfer the media into sterile petri dishes (25 cc per dish) via sterile pipette in a laminar flow hood. Once the media cools (partially covered) to room temperature and solidifies and once any condensation on the petri dishes had dissipated, transfer the plates to sterile packages and stored at 4 °C.

Loading Dye 2x

Ingredient	Concentration
TRIS HCl	0.09 M
Glycerol	20%
SDS	10 G
dd H_2O	qs 1 L

Do not diule prior to use.

Lysing Buffer

Ingredient	Quantity
TRIS	0.242 G
EDTA 0.5 M	160 μl
DTT	30.8 mg
Lysozyme	8 mg
dd H_2O	qs 40 ml

Adjusted the pH to 9.0 with 1 M HCl and trickle argon gas through the buffer for 5 minutes before adding DTT.

Minimal Media

Part A - M9 salts

Ingredient	Quantity
Na_2HPO_4	6.78 G
KH_2PO_4	3.0 G
NaCl	500 mg
NH_4Cl	1 G
dd H_2O	946 ml

Part B

Ingredient	Quantity
Glucose	4 G
Thiamine 1%	1 ml
$FeSO_4$ 1 mM	2 ml
$MgSO_4$ 1 M	2 ml
$CaCl_2$ 1 M	100 μl
dd H_2O	24 ml

Combine **Part A** ingredients and autoclave at 121 °C and 15 psi for 30 minutes. Once **Part A** has cooled to 25 °C, add the ingredients in **Part B** and with the appropriate quantity of antibiotics (ampicillin 100 μg/ml, chloramphenicol 50 μg/ml) via filter sterilization through a 22 μm pore filter.

Phosphate Buffered Saline

Ingredient	Quantity
NaCl	8 G
KCl	200 mg
Na_2HPO_4	1.44 G
KH_2PO_4	240 mg
dd H_2O	800 ml

Dissolve ingredients and adjust the pH to 7.4 with HCl.

SDS PAGE Gel

Ingredient	Quantity	
	Main 12%	**Stacking 4%**
dd H_2O	3.4 ml	3.1 ml
Acrylamide 30%	4.0 ml	670 μl
TRIS HCl pH 8.8 1.5 M	2.5 ml	-
TRIS HCl pH 6.8 0.5 M	-	1.25 ml
SDS 20%	50μl	25 μl
APS 10%	50μl	25 μl
TEMED	6μl	5 μl

Prepared the gels with fresh APS. Add the ingredients in the order indicated, such that the catalyst, TEMED, is added last just prior to pouring the unset gel into molds. Fill the molds $\sim 3/4$ full with the main gel and layer with n-butanol. Once set (~20 minutes), layer the stacking gel and insert a lane divider.

Teriffic Broth Media

Part A

Ingredient	Quantity
Tryptone-Peptone	12 G
Yeast Extract	24 G
Glycerol	4 ml
dd H_2O	900 ml

Part B

Ingredient	Quantity
KH_2PO_4	2.31 G
K_2HPO_4	15.54 G
dd H_2O	90 ml

Make Part B up to 100 ml. Autoclave the parts separately at 15 psi and 121 °C for 30 minutes. Once cooled to room temperature, combined the parts and add the antibiotics *via* filter sterilization (final concentrations: 100 μg/ml for ampicillin and 50 μg/ml for chloramphenicol).

Appendix B

HRV14-3C Protease Expression and Purification Protocols

B.1 Expression Protocol

Plasmid vector: pET-3a containing the HRV14-3ABC gene
Host cells: E. *coli* BL21(DE3) pLysS cells
Freezer stock: Oct22-D

1. Streak plates with OCT22-D 50% freezer stock of E. *coli* (BL21(DE3) pLysS), which contains the pET-3a vector with the HRV14-3ABC gene, onto LB media plates containing 100 μg/ml of ampicillin and 50 μg/ml of chloramphenicol.

2. Incubate plates at 37 °C for 24 hours.

3. Inoculate 25 cc of TB or LB media with a few isolated colonies.

4. Grow cells at 37 °C for 12 to 24 hours[1] oh a rotobed shaker (325 RPM). Give booster doses of ampicillin (100 μg/ml) and chloramphenicol (50 μg/ml) at 6 hour intervals to maintain plasmid presence.

5. Transfer sufficient amount of cells to 1 L TB media to achieve an initial OD_{600nm} $\sim$0.1[2] and divide into four 1 L beveled flasks (250 cc/flask).

6. Grow cells at 37 °C on rotobed shaker (325 RPM) until OD_{600nm} $\sim$1.0

[1] Requires 12 hours in TB or LB media; 24 hours in MM
[2] Requires $\sim$1% or 2.5% v/v of innoculum for TB/LB or MM overnight cultures respectively

7. Centrifuge cells at 2300 x g for 30 minutes. Decant media and re-suspend cells in 250 cc of labeled MM containing (100 μg/ml) and chloramphenicol (50 μg/ml). Transfer cells to 1 L bevelled flask and return to rotobed shaker (325 RPM) at 37 °C.

8. Following a 1 hour recovery period, induce cells with 0.04% of IPTG[3]. Reduce temperature to 25 - 30 °C and allow expression to continue for 12 to 14 hours. Give booster doses of 100 μg/ml of ampicillin after 6 hours of induction to maintain pET-3a plasmid.

9. Harvest cells by centrifugation (2300 x g) for 30 minutes. Decant media, wash with PBS and re-centrifuge[4]. Decant PBS and store whole cells at -20 °C until required for purification.

B.2 Protein Induction Test

Prior to proceeding with protein purification, protease expression was confirmed by running an aliquot of production culture on 12% SDS-PAGE gel for analysis[5].

1. Centrifuge 500μl aliquot at 15,000 RPM in a 1.5 cc epindorph tube for 15 minutes. Decant PBS. Store cells at -20 °C till required
2. Resuspend cells in 100 μl of dd H_2O and vortex.
3. Combine with 100 μl of 2x loading dye and boil for 10 minutes.
4. Separate 2.5μl, 5μl and 10μl loads of aliquot along with standard in 12% SDS-PAGE gel. Compare against standard to confirm induction.

[3]Requires $\sim$ 500μl of a 20% IPTG solution in 250 ml of culture

[4]Retain 500 μl of cells for induction test

[5]Cell lysis can be started during the induction test, which should be completed before lysed cells are ready for loading onto the chromatography column

B.3 Protein Purification

1. Re-suspend cells in 40 cc lysing buffer. Let cells equilibrate to room temperature for 20 minutes.
2. Transfer lysing buffer with suspended cells (4 x 10 cc) to four 25 cc Nalgene® centrifuge tubes.
3. Complete 3 cycles of freezing at -70 °C for 20 minutes and thawing at 4 °C for 30 minutes.
4. Let lysed cells stand additional 30 minutes following freezing and thawing cycles.
5. Pellet cellular material at 18,000 RPM (~36,200 x g) for 60 minutes. Decant supernatant, add ~8 - 10 drops of PEE / tube, lightly agitate and re-pellet at 18,000 RPM (~36,200 x g) for 20 minutes.
6. Gently recover clear supernatant and adjust pH to 9.0 with TRIS.
7. Layer supernatant over either a Q-Sepharose™or DEAE-Sepharose™anionic exchange chromatography column[6] (~120 cc) pre-equilibrated to pH 9.0 at 4 °C with chromatography buffer[7]. Run column at a rate of 2.5 - 3 cc/min. and collect 9 ml fractions[8].
8. Pool fractions containing HRV14-3C and dialyzed in phosphate buffer using 5000 MW cut-off dialysis membranes[9]. Use the Bradford assay to quantify the purified HRV14-3C protease.

Addendum: Following step 7 and prior to dialysis, the use of a hypoxyapitite column was used to improve the purification and stability of the apo HRV14-3C protease. A detailed summary of its use follows:

> Collect fractions containing pure HRV14-3C protease[10] and load onto a 50 ml BioGel-HA® hypoxyapatite column. Wash column with 100 ml of loading buffer (50 mM TRIS, 1 mM EDTA, 5 mM DTT, pH 8.5) and 100 ml of phosphate buffer (40 mM KH_2PO_4, 0.5 mM EDTA, 10 mM DTT, 1 mM NaN_3, pH 6.5). Elute the HRV14-3C protease with a 40 to 195 mM phosphate buffer linear gradient using a flow rate of 2 to 3 ml/min.

[6]Clean and sterilzize chromatorgraphy columns (1.5 cm x 30 cm) 95% ethanol. Plug with wet glass wool and fill with 120 to 150 ml of ion exchange resin (follow manufacturer's recomendations for resuspension and packing. Degassed all aggiated resins under vacuum prior to use). Packing of the resin should be done with a flow rate of 2.5 cc/min using 3 to 4 volumes of chromatography buffer at 4 °C prior to use

[7]The pH of TRIS buffer is inversely related to temperature: pH = pH - 0.03 (Δ °C). Therefore, temperature regulation can affect protein separation

[8]HRV14-3C should elude just following the initial flow-through as detected with UV_{280nm} absorption

[9]Prepare dialysis membranes by pre-soaking in dd H_2O for 30 minutes, 20 mM $NaHCO_2$ for 30 minutes then 0.5 mM EDTA for 30 minutes and checked for leaks

[10]Confirm *via* pNA assay and SDS-PAGE gel analysis

Appendix C

XPLOR / CNS Ace-LEALFQ-ethylpropionate inhibitor input scripts

The following files were written to incorporate the ethylpropionate group into the XPLOR and CNS structure calculations. The *topology* file defines the atoms, bonds and angles in the ethylpropionate group and futhermore, includes all the patches called by the *annealing.inp* script. This file is required to build the ethylpropionate ligand and covalently attach it to the extended chain HRV14-3C protease prior to running the simulated annealing protocol. The *parameter* file contains the specific lengths, angles and respective energy scalars for the atoms and angles defined in the *topology* file. These files need to be appended to the **.top* and **.par* files used or called seperately by the *generate.inp*, *generate_extended.inp*, and *annealing.inp* scripts during the structure calculation. Additionally, modifications to the RECOORD protocols were also required. All these changes have been itemized in this appendix and electronic copies of these files have been deposited into the PDB along with the inhibited HRV14-3C protease structures (accession code 2B0F.pdb)

C.1 Ethylpropionate *topology* File

```
!
! Topology file for incorporation of the ethylpropionate moiety
! into XPLOR/CNS calcuations
!
! Written by Trent C. Bjorndahl, July 25, 2005
! Final modifications made Nov. 15, 2005 to correct angle strain about the covalent bond,
! which was fixed with the inclusion of the 'THRE' patch for a novel three residue
! angle parameter
!


RESIdue CAP ! Defines the atoms, their charges and bonds for the ethylpropionate group
GROUp
     ATOM CX5  TYPE    CH1E     CHARge   0.0    END
     ATOM CX6  TYPE    CH2E     CHARge   -0.2   END
     ATOM CX7  TYPE    C        CHARge   -0.5   END
     ATOM OX8  TYPE    O        CHARge   -0.62  END
     ATOM OX9  TYPE    OC       CHARge   -0.62  END
     ATOM CJ1  TYPE    CH2E     CHARge   -0.3   END
     ATOM CJ2  TYPE    CH3E     CHARge   0.0    END
     ATOM H1   TYPE    HA       CHARge   0.1    END
     ATOM H2   TYPE    HA       CHARge   0.1    END
     ATOM H3   TYPE    HA       CHARge   0.1    END
     ATOM H4   TYPE    HA       CHARge   0.1    END
     ATOM H5   TYPE    HA       CHARge   0.1    END
     ATOM H6   TYPE    HA       CHARge   0.1    END
     ATOM H7   TYPE    HA       CHARge   0.1    END
     ATOM H8   TYPE    HA       CHARge   0.1    END


     BOND CX5 CX6     BOND CX6 CX7     BOND CX7 OX8
     BOND CX7 OX9     BOND OX9 CJ1     BOND CJ1 CJ2

     BOND CX5 H1
     BOND CX6 H2
     BOND CX6 H3
     BOND CJ1 H4
     BOND CJ1 H5
     BOND CJ2 H6
     BOND CJ2 H7
     BOND CJ2 H8

     IMPRoper CX5 CX6 CX7 OX8
     IMPRoper CX5 CX6 CX7 OX9
     IMPRoper CX6 CX7 OX9 CJ1
     IMPRoper OX8 CX7 OX9 CJ1
     IMPRoper CX7 OX9 CJ1 CJ2

     IMPRoper CJ1 OX9 CX7 OX8 ! chirality or flatness improper
     IMPRoper H6 H7 CJ1 H8 ! methyl improper

     ACCEptor OX8 " "
     ACCEptor OX9 " "
END { RESIdue CAP }
```

```
PRESidue COVA ! adds the covalent bond between Cys146 SG and the ethylpropionate CX5 carbon
     DELEte DONOR -HG -SG
     DELEte atom -HG end
     ADD BOND -SG +CX5
     MODIfy ATOM -SG TYPE=SM charge=-0.470 end

     ADD ANGLe -CB -SG +CX5
     ADD ANGLe -SG +CX5 +H1
     ADD ANGLe -SG +CX5 +CX6
END { COVA }

PRESidue THRE ! for 3 residue link
     ANGLe +SG -CX5 -CA
END { THRE }

PRESidue PRET ! Patch added for ethylpropionaote group by TCB
     DELEte atom -C end
     DELEte atom -O end

     ADD BOND -CA +CX5
     ADD ANGLe +H1 +CX5 -CA
END { PRET }

PRES NACE ! Added by TCB to add N-term Acetyl group
     ADD BOND -C +N

     ADD ANGLe -CA -C +N
     ADD ANGLe -O -C +N
     ADD ANGLe -C +N +CA
     ADD ANGLe -C +N +HN

     ADD DIHEdral -C +N +CA +C mult 6
     ADD DIHEdral -C +N +CA +CB mult 6

     ADD IMPRoper -CA -C +N +CA ! omega angle across peptide plane
END { NACE }

PRES NACA
     !! dummy patch for first in *.pep file
END { NACA }

PRES CAPE
     !! dummy patch for last in *.pep file
END { CAPE }
```

C.2 Ethylpropionate *parameter* File

```
!
! Bond length, dihedral and improper angles for the ethylpropionate moiety
! Created by Trent Bjorndahl
! Final revisions: Nov. 25, 2005
!

set echo=false end

BOND C    CH3E  1000.000  1.516 ! from C to CH2E
BOND CT   CH1E  1000.000  1.530 !
BOND SM   CH1E  1000.000  1.800 ! taken from Met
BOND OC   CH2E  1000.000  1.530 !
BOND CT   CT    1000.000  1.530 !
BOND CT   C     1000.000  1.535 !
BOND CT   HA    1000.000  1.080 !


!acetyl cap
ANGLe C     CH3E  HA    500.0   109.50
ANGLe CH3E  C     O     500.0   120.8258
ANGLe CH3E  C     NH1   500.0   116.1998


!covalent bond
ANGLe CH2E  SM    CH1E  500.0   100.8987 ! taken from met
ANGLe SM    CH1E  HA    500.0   108.6768 ! taken from met
ANGLe SM    CH2E  CH1E  500.0   112.6822 ! taken from met
ANGLe SM    CH1E  CH2E  500.0   112.6822 ! taken from met
ANGLe SM    CH1E  CH1E  500.0   111.3300 ! three residue angle


!ester group
ANGLe OC    C     O     500.0   123.3548
ANGLe C     OC    CH2E  500.0   124.0000
ANGLe OC    CH2E  CH3E  500.0   120.8258
ANGLe OC    CH2E  HA    500.0   109.5000


! covalent bond improper
IMPRoper HA    HA    CH1E   SM    500.00 {sd=0.031}  0   -72.4655

! impropers for ester planarity
IMPRoper C     OC    CH2E   CH3E  500.0 {sd=0.031}   0   180.00
IMPRoper O     C     OC     CH2E  500.0 {sd=0.031}   0     0.00
IMPRoper CH2E  C     OC     CH2E  500.0 {sd=0.031}   0   180.00
IMPRoper CH1E  CH2E  C      O     500.0 {sd=0.031}   0     0.00
IMPRoper CH1E  CH2E  C      OC    500.0 {sd=0.031}   0   180.00

! methyl improper
IMPRoper HA    HA    C      HA    500.0              0   -66.5934
IMPRoper HA    CH2E  HA     HA    500.0              0   -66.5934

! acetyl dihedrals angle
IMPRoper CH3E  C     NH1    CH1E  500.0              0   180.00
IMPRoper C     CH3E  NH1    O     500.0              0     0.00


!dihedrals
DIHEdral NH1   CH1E   CH2E   SM   2.00              3     0.0

set echo=true end
```

C.3 XPLOR/CNS *generate.inp* File Modifications

```
!
! The following additions need to be inserted into the generate.inp script
! following the "segment" generation section
! The generate.inp script is used by XPLOR/CNS to generate the extended protein
! These changes call the patches added to the *.top file
! written by Trent C. Bjorndahl
! final revision, Nov. 15, 2005
!

segment name="HRV "
chain
if ( convert = true ) then
convert = true
end if
if ( separate = true ) then
separate = true
end if
 topallhdg5.3.pep



sequence
gly  pro  asn  thr  glu  phe  ala  leu  ser  leu
leu  arg  lys  asn  ile  met  thr  ile  thr  thr
ser  lys  gly  glu  phe  thr  gly  leu  gly  ile
his  asp  arg  val  cys  val  ile  pro  thr  his
ala  gln  pro  gly  asp  asp  val  leu  val  asn
gly  gln  lys  ile  arg  val  lys  asp  lys  tyr
lys  leu  val  asp  pro  glu  asn  ile  asn  leu
glu  leu  thr  val  leu  thr  leu  asp  arg  asn
glu  lys  phe  arg  asp  ile  arg  gly  phe  ile
ser  glu  asp  leu  glu  gly  val  asp  ala  thr
leu  val  val  his  ser  asn  asn  phe  thr  asn
thr  ile  leu  glu  val  gly  pro  val  thr  met
ala  gly  leu  ile  asn  leu  ser  ser  thr  pro
thr  asn  arg  met  ile  arg  tyr  asp  tyr  ala
thr  lys  thr  gly  gln  cys  gly  gly  val  leu
cys  ala  thr  gly  lys  ile  phe  gly  ile  his
val  gly  gly  asn  gly  arg  gln  gly  phe  ser
ala  gln  leu  lys  lys  gln  tyr  phe  val  glu
lys  gln  end
     end
end


segment name="INH "
chain
@topallhdg5.3.pep
sequence
ace leu glu ala leu phe gln cap end
end
end

patch COVA
reference=" - "=( resid 146 and segid HRV )
reference=" + "=( resid 8 and segid INH )
end

patch THRE
reference=" + "=(resid 146 and segid HRV )
reference=" - "=(resid 8 and segid INH )
reference=" - "=(resid 7 and segid INH )
end
```

C.4 Peptide Linkage File

```
SET ECHO=FALSE END

!
! Modified from 'topallhdg5.3.pep '
! Generates protein peptide bonds and termini for a protein seq
! Modified by Trent C. Bjorndahl, Nov. 15, 2005
! to incorporate the acetyl N-terminal acetyl and C-terminal ethylpropionate groups
! It should be called from the 'generate.inp ' in the SEGMent SEQUence level
!


link PPGP  head - GLY  tail + PRO end
link PPGG  head - GLY  tail + GLY end
link PEPP  head - *    tail + PRO end
link PPG2  head - GLY  tail + *   end
link PPG1  head - *    tail + GLY end
link PEPT  head - *    tail + *   end

! the following linkages added by TCB for ethyl propenoate group

link PRET  head - GLN  tail + CAP end { LINK to ethylproeonate group }
llink NACE  head - ACE  tail + LEU end { LINK to N-term acetyl group }

first NTER             tail + GLY end ! modified by TCB for HRV14-3C sequence only
first NACA             tail + ACE end

last CTER head - GLN             end ! modified by TCB for HRV14-3C sequence only
last CAPE head - CAP             end


SET ECHO=TRUE END
```

C.5 RECOORD Input File Modifications

Specific changes need to be made to the RECOORD water refinement protocols in order to incorpoate the ethylpropionate inhibitor into the inhibited HRV14-3C structure calculations. Generally, the files needed for water refinement are created by the input shell scripts (*annealing.sh and re_h2o.sh*). However, because the ligand is novel and includes "non-standard", some files need ammendments and confirmation of modification. The specific changes include:

1. *annealing.inp*
   select torsion protocol

2. *re_h2o.inp*
   select the "*OPLSX*" non-bonded parameter

3. *read_data.cns*
   noe
   averaging * cent ! was sum

4. *generate.inp*
   If the inhibitor *parameter and topology* files are not appended to toppallhdg5.3.* files, but called seperately, the directories they reside in need to be set here:

   {==== ligand topology and parameter files ====}
   {* ligand topology file *}
   {===>} ligand_topology_infile= "SCRIPTS:toppar/eth.top";
   {* ligand parameter file *}
   {===>} ligand_parameter_infile= "SCRIPTS:toppar/eth.par";

   choose the OPLSX or PROLSQ non-bonded parameter set for water refinement:
   {===>} evaluate ($par_nonbonded= "OPLSX")

5. *Run.cns*[1]
   confirm the segment ID codes for the protein and inhibitor

   {==== filenames ====}
   {* the name of your current project *}
   {===>} fileroot= "rootStructure";
   {===>} prot_segid_A= "HRV ";
   {===>} prot_segid_B= "INH ";

   and confirm the non-bonded parameter set selection
   {* type of non-bonded parameters *}
   {===>} par_nonbonded = "OPLSX";

[1]This file should be created automatically and put in the working project directory

Appendix D

Chemical Shift Assignment Tables

The stereospecific labels in these tables are for annotation purposes only. No experiments were preformed to determine the stereospecificity of the branch chain amino acid methyl groups.

D.1 Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{13}C Chemical Shifts

Table D.1: Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{13}C Chemical Shift Assignments†

Group	$^{13}C_\alpha$	$^{13}C_\beta$	$^{13}C_\gamma$	Other
Acetyl				44.9
P_6 Leu	55.6	44.7	28.1	$^{13}C_{\delta 1}$ 25.8; $^{13}C_{\delta 2}$ 25.8
P_5 Glu	53.3	33.4	35.7	
P_4 Ala	52.7	21.8		
P_3 Leu	55.5	44.5	28.3	$^{13}C_{\delta 1}$ 27.3; $^{13}C_{\delta 2}$ 27.1
P_2 Phe	58.3	41.8		$^{13}C_\delta$ 133.2; $^{13}C_\epsilon$ 130.7; $^{13}C_{\delta\zeta}$, 132.8
P_1 Gln *trans(E)*	56.5	31.1	34.4	
P_1 Gln *cis(Z)*	51.2	31.1	34.4	
H_1 & H_2 *trans(E)*				153.3, 124.4
H_1 & H_2 *cis(Z)*				153.6, 123.6
Ethyl CH_2				64.2
Ethyl CH_3				18.3

†Values recorded in DMSO

D.2 Acetyl-LEALFQ-ethyl Propenoate Inhibitor ^{1}H Chemical Shifts

Table D.2: **Acetyl-LEALFQ-ethylpropenoate Inhibitor ^{1}H Chemical Shift Assignments**[†‡].

Group	^1H_N	$^1H_\alpha$	$^1H_\beta$	$^1H_\gamma$	Other
Acetyl					1.63 (57)
P_6 Leu	7.04 (50)	3.37 (49)	0.48 (54), 0.64 (54)	0.63 (55)	$^1H_{\delta 1}$ -0.04 (56); $^1H_{\delta 2}$ -0.09 (58)
P_5 Glu	7.10 (39)	3.49 (38)	0.78 (45), 0.87 (45)	1.18 (46)	
P_4 Ala	7.03 (28)	3.33 (27)	0.29 (37)		
P_3 Leu	7.06 (14)	3.33 (14)	0.55 (24), 0.72 (24)	0.72 (25)	$^1H_{\delta 1}$ 0.00 (26); $^1H_{\delta 2}$ -0.04 (36)
P_2 Phe *trans(E)*	6.95 (9)	3.59 (8)	1.97, 2.10 (7)		$^1H_\delta$ 6.24 (2,6); $^1H_\epsilon$ 6.32 (3,5); $^1H_\zeta$, 6.43 (4)
P_2 Phe *cis(Z)*	6.90 (9)	3.48 (8)	1.97, 2.10 (7)		$^1H_\delta$ 6.24 (2,6); $^1H_\epsilon$ 6.32 (3,5); $^1H_\zeta$, 6.43 (4)
P_1 Gln *trans(E)*	7.14 (18)	3.33 (19)	0.87 (20), 1.02 (20)	1.36 (21)	$^1H_\epsilon$ 5.88, 6.35 (35)
P_1 Gln *cis(Z)*	7.05 (18)	4.37 (19)	0.87 (20), 1.02 (20)	1.36 (21)	$^1H_\epsilon$ 5.07, 5.65 (35)
H_1 & H_2 *trans(E)*					5.81 (32), 4.78 (31)
H_1 & H_2 *cis(Z)*					5.07 (32), 4.92 (31)
Ethyl CH_2					3.24 (43)
Ethyl CH_3					0.35 (44)

[†]Values recored in DMSO

[‡]Numbers in parenthesis correspond to labels given to protons in Figure 3.6

D.3 Apo HRV14-3C Chemical Shifts

Table D.3: **Chemical Shift Assignments for the Apo HRV14-3C protease**[1]

Residue	N	C′	C_α	C_β	Other
P^{2}		177.3	64.6 (4.39)	32.0 (2.45, 2.45)	C_γ, 27.1 (1.94, 1.94); C_δ, 49.5 (3.55, 3.65)
N^{3}	115.7 (8.91)	176.5	53.5 (5.01)	39.0 (2.77, 3.00)	$N_{\delta 2}$, 113.0 (6.88, 7.59)
T^{4}	116.9 (7.70)	174.8	67.2 (3.75)	68.7 (4.17)	$C_{\gamma 2}$, 21.9 (1.17)
E^{5}	121.2 (8.61)	179.2	60.3 (3.91)	28.8 (2.00, 2.05)	C_γ, 36.4 (2.28, 2.33)
F^{6}	121.2 (8.60)	177.4	61.4 (4.21)	39.6 (3.30, 2.81)	C_δ, * (7.03); C_ϵ, * (6.54); C_ζ, * (*)
A^{7}	122.7 (8.32)	179.1	55.6 (3.70)	18.7 (1.32)	
L^{8}	117.0 (8.60)	179.6	57.7 (3.96)	41.5 (1.32, 1.83)	C_γ, 26.9 (1.66); $C_{\delta 1}$, 25.5 (0.77); $C_{\delta 2}$, 22.9 (0.77)
S^{9}	115.5 (7.79)	177.1	61.3 (4.15)	62.8 (3.92, 3.92)	
L^{10}	121.3 (7.68)	*	57.9 (3.78)	41.9 (1.08, 1.71)	C_γ, 25.9 (1.14); $C_{\delta 1}$, 23.3 (0.47); $C_{\delta 2}$, * (0.53)
L^{11}	120.2 (8.15)	177.7	57.9 (3.84)	41.5 (1.65, *)	C_γ, 27.3 (1.38); $C_{\delta 1}$, 23.5 (0.67); $C_{\delta 2}$, 24.5 (0.71)
R^{12}	115.2 (7.75)	178.5	58.7 (4.12)	30.9 (1.85, *)	C_γ, 27.9 (1.75, 1.65); C_δ, 43.3 (3.14, *)
K^{13}	113.8 (7.69)	177.8	57.2 (4.69)	35.3 (*)	C_γ, 25.7 (1.49, *); C_δ, 29.0 (*, *); C_ϵ, 42.1 (2.94, 2.94)
N^{14}	116.7 (8.63)	172.9	56.1 (5.05)	43.5 (2.08, 3.15)	$N_{\delta 2}$, 112.4 (6.95, *)
I^{15}	121.3 (7.43)	174.3	60.9 (5.04)	39.1 (1.81)	$C_{\gamma 1}$, 28.9 (0.99, 1.35); $C_{\gamma 2}$, 18.9 (0.99); $C_{\delta 1}$, 14.2 (0.64)
M^{16}	124.2 (8.64)	174.3	53.7 (4.92)	37.6 (1.80, 2.11)	C_γ, 30.9 (2.30, 2.37)
T^{17}	118.3 (8.86)	174.9	62.9 (5.05)	69.2 (3.98)	$C_{\gamma 2}$, 23.04 (1.08)
I^{18}	131.1 (9.73)	175.1	58.5 (5.49)	38.7 (1.77)	$C_{\gamma 1}$, 27.5 (1.26, 1.47); $C_{\gamma 2}$, 16.6 (0.79); $C_{\delta 1}$, 12.7 (0.77)
T^{19}	124.5 (9.44)	176.0	61.8 (5.52)	69.9 (4.01)	$C_{\gamma 2}$, 21.9 (1.33)
T^{20}	120.1 (9.39)	*	59.3 (5.33)	71.2 (4.78)	$C_{\gamma 2}$, 22.0 (0.98)
S^{21}	* (*)	175.3	61.5 (4.42)	62.8 (4.08, 4.08)	
K^{22}	118.5 (8.41)	176.0	54.5 (4.53)	32.1 (2.28, 2.28)	C_γ, 24.6 (1.42, 1.63); C_δ, 27.5 (1.76, *); C_ϵ, 42.1 (2.97, *)
G^{23}	108.2 (7.67)	171.2	44.1 (4.49, 3.78)		
E^{24}	117.7 (8.23)	176.3	55.9 (4.94)	31.9 (1.59, 1.84)	C_γ, 36.4 (2.15, 2.24)
F^{25}	120.3 (9.43)	175.6	57.1 (4.50)	43.0 (2.65, 2.92)	C_δ, * (7.25); C_ϵ, * (7.11); C_ζ, * (*)
T^{26}	120.2 (8.43)	172.5	65.2 (4.74)	69.8 (3.86)	$C_{\gamma 2}$, 22.9 (1.34)
G^{27}	111.0 (9.10)	171.3	44.0 (5.04, 3.21)		
L^{28}	121.0 (7.59)	175.4	52.7 (4.95)	44.2 (1.20, 1.79)	C_γ, 26.6 (1.35); $C_{\delta 1}$, 22.9 (0.80); $C_{\delta 2}$, 26.6 (0.65)
G^{29}	117.0 (9.29)	172.3	43.3 (4.15, 3.16)		
I^{30}	126.9 (8.80)	173.8	59.9 (3.91)	37.0 (1.55)	$C_{\gamma 1}$, 25.2 (1.40, 0.70); $C_{\gamma 2}$, 17.2 (0.13); $C_{\delta 1}$, 8.74 (0.20)
H^{31}	108.6 (7.34)	172.0	54.7 (4.44)	30.1 (3.02, 3.74)	
D^{32}	122.9 (10.13)	174.2	57.4 (4.03)	40.8 (2.61, 2.61)	
R^{33}	125.8 (8.34)	175.1	56.4 (4.73)	29.1 (2.06, *)	C_γ, 29.7 (1.53, *); C_δ, 43.7 (3.09, 3.09)
V^{34}	122.0 (8.99)	175.4	62.8 (4.99)	33.3 (1.95)	$C_{\gamma 1}$, 21.7 (0.98); $C_{\gamma 2}$, 21.7 (0.98)
C^{35}	124.0 (9.65)	*	55.0 (5.53)	32.7 (2.47, 2.58)	
V^{36}	114.9 (8.28)	173.3	59.0 (5.55)	36.3 (2.30)	$C_{\gamma 1}$, 22.0 (1.11); $C_{\gamma 2}$, 22.6 (1.11)

Continued on Next Page...

Table D.3: Apo HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
I^{37}	116.4 (8.29)		58.2 (4.98)	43.6 (1.55)	$C_{\gamma 1}$, 24.0 (0.80, 1.30); $C_{\gamma 2}$, 17.9 (0.79); $C_{\delta 1}$, 14.9 (0.48)
P^{38}		175.3	63.3 (4.36)	32.1 (1.93, 2.37)	C_γ, 27.4 (2.01, *); C_δ, 51.1 (3.61, 3.67)
T^{39}	130.3 (7.95)	*	61.7	67.2	$C_{\gamma 2}$, 21.5 (1.14)
H^{40}	* (*)	174.8	58.4 (4.22)	27.6 (2.85, 2.94)	$C_{\epsilon 1}$, * (7.20)
A^{41}	121.0 (7.20)	177.4	54.1 (3.42)	18.0 (0.08)	
Q^{42}	110.3 (8.04)		56.2 (4.01)	28.0 (1.63, 1.80)	C_γ, 34.1 (2.09, 2.30); $N_{\epsilon 2}$, 111.3 (6.85, 7.38)
P^{43}		176.9	62.9 (4.27)	31.9 (1.64, 1.83)	C_γ, 28.0 (1.73, *); C_δ, 49.8 (2.93, 3.87)
G^{44}	112.7 (9.18)	173.2	44.0 (3.73, 4.39)		
D^{45}	121.3 (8.43)	175.4	56.2 (4.54)	41.1 (2.48, 2.86)	
D^{46}	114.9 (7.82)	173.8	52.7 (5.19)	44.4 (2.33, 2.56)	
V^{47}	111.7 (8.80)	174.7	59.1 (4.72)	35.9 (2.07)	$C_{\gamma 1}$, 22.3 (0.70); $C_{\gamma 2}$, 17.9 (0.50)
L^{48}	116.1 (7.96)	177.3	53.4 (5.29)	42.4 (0.88, 1.93)	C_γ, 26.5 (1.57); $C_{\delta 1}$, 22.0 (0.66); $C_{\delta 2}$, 24.9 (0.83)
V^{49}	122.8 (9.41)	176.5	61.3 (4.58)	32.3 (2.04)	$C_{\gamma 1}$, 21.6 (0.81); $C_{\gamma 2}$, 21.6 (0.81)
N^{50}	129.0 (9.87)	175.7	54.2 (4.49)	37.0 (3.03, 3.17)	
G^{51}	101.3 (8.35)	173.2	45.7 (3.52, 4.34)		
Q^{52}	120.8 (7.81)	175.2	54.1 (4.61)	30.4 (2.01, 2.12)	C_γ, 33.5 (2.27, 2.37)
K^{53}	127.9 (8.88)	175.5	58.3 (4.46)	32.1 (1.52, 1.77)	C_γ, 24.8 (1.21, *); C_δ, 29.3 (1.57, *); C_ϵ, 41.7 (2.82, 2.82)
I^{54}	131.0 (9.39)	174.4	58.4 (4.50)	40.5 (1.77)	$C_{\gamma 1}$, 27.1 (1.29, 1.56); $C_{\gamma 2}$, 18.8 (0.93); $C_{\delta 1}$, 10.9 (0.81)
R^{55}	126.2 (8.90)	175.0	56.9 (4.51)	31.1 (1.97, 2.12)	C_γ, 27.8 (1.57, 1.80); C_δ, 43.5 (3.29, 3.34)
V^{56}	125.3 (8.44)	175.9	62.2 (4.03)	32.2 (2.03)	$C_{\gamma 1}$, 22.3 (0.80); $C_{\gamma 2}$, 23.3 (0.80)
K^{57}	129.1 (8.88)	175.9	58.0 (4.19)	33.3 (1.45, 1.69)	C_γ, 24.9 (1.39, *); C_δ, 29.15 (1.58, *); C_ϵ, 41.7 (2.90, 2.90)
D^{58}	114.8 (7.67)	173.0	53.5 (4.82)	44.4 (2.53, 2.67)	
K^{59}	118.1 (8.20)	174.4	55.0 (5.39)	35.4 (1.40, 1.72)	C_γ, 22.9 (1.20, 1.29); C_δ, 29.9 (1.29, 1.49); C_ϵ, 41.8 (2.86, 2.93)
Y^{60}	119.2 (8.45)	*	56.6 (4.77)	40.1 (2.92, 3.14)	C_δ, * (6.92); C_ϵ, * (6.72); C_ζ, * (*)
K^{61}	123.2 (8.58)	*	55.6 (4.61)	33.0 (1.73, *)	C_γ, 24.6 (1.37, *); C_δ, 29.0 (*, *); C_ϵ, 42.4 (2.91, *)
L^{62}	* (*)	*	55.4 (4.44)	* (1.00, 1.56)	C_γ, 27.2 (1.30); $C_{\delta 1}$, 24.5 (0.29); $C_{\delta 2}$, 23.2 (0.53)
V^{63}	* (*)	175.0	59.6 (4.65)	35.8 (1.80)	$C_{\gamma 1}$, 18.8 (0.93); $C_{\gamma 2}$, 21.7 (0.82)
D^{64}	126.0 (8.94)		53.0 (4.75)	41.5 (2.56, 3.35)	
P^{65}		177.4	65.5 (4.41)	31.7 (2.38, 2.48)	C_γ, 28.0 (1.95, 2.12); C_δ, 51.6 (3.65, 3.56)
E^{66}	116.8 (8.28)	*	55.8 (4.37)	29.5 (2.28, *)	C_γ, 37.4 (2.18, *)
N^{67}	* (8.62)	174.7	54.5 (4.21)	36.4 (2.98, 3.11)	$N_{\delta 2}$, 112.1 (6.92, 7.56)
I^{68}	121.6 (8.47)	175.8	61.1 (4.15)	38.7 (1.85)	$C_{\gamma 1}$, 27.2 (1.16, 1.43); $C_{\gamma 2}$, 17.6 (0.85); $C_{\delta 1}$, 13.0 (0.77)
N^{69}	124.7 (8.43)	*	56.6 (4.25)	* (2.90, *)	$N_{\delta 2}$, * (*, *)
E^{71}	* (*)	*	* (4.15)	* (2.06, *)	C_γ, 38.3 (2.31, 2.43)
L^{72}	* (*)	175.2	53.2 (4.81)	46.8 (1.07, 1.40)	C_γ, 26.3 (0.95); $C_{\delta 1}$, 23.5 (0.81); $C_{\delta 2}$, 23.5 (0.81)
T^{73}	118.8 (8.75)	171.4	62.1 (4.89)	72.3 (3.57)	$C_{\gamma 2}$, 20.7 (0.84)
V^{74}	127.4 (8.75)	175.5	61.3 (4.53)	32.6 (1.39)	$C_{\gamma 1}$, 22.4 (0.98); $C_{\gamma 2}$, 21.4 (0.78)
L^{75}	126.3 (9.34)	175.8	53.4 (5.05)	44.8 (1.28, 1.95)	C_γ, 26.6 (1.55); $C_{\delta 1}$, 25.7 (0.68); $C_{\delta 2}$, 24.8 (0.68)
T^{76}	117.5 (8.66)	175.0	61.9 (5.05)	69.0 (3.97)	$C_{\gamma 2}$, 21.7 (1.16)

Continued on Next Page...

Table D.3: Apo HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
L^{77}	126.4 (9.24)	176.2	54.0 (4.62)	44.4 (1.21, 1.81)	C_γ, 26.9 (1.82); $C_{\delta 1}$, 21.8 (0.61); $C_{\delta 2}$, 26.9 (0.90)
D^{78}	121.4 (8.59)	173.5	52.5 (4.81)	39.8 (2.05, 3.17)	
R^{79}	119.3 (7.28)	*	54.2 (4.73)	33.6 (*)	C_γ, * (*, *); C_δ, * (*, *)
N^{80}	* (*)	176.3	55.6 (4.45)	39.2 (2.70, 3.10)	$N_{\delta 2}$, * (*)
E^{81}	120.1 (8.36)	175.3	57.6 (3.98)	30.5 (1.61, 1.67)	C_γ, 36.6 (2.15, *)
K^{82}	121.4 (8.38)	177.5	54.9 (4.17)	35.4 (1.64, 1.89)	C_γ, 25.8 (1.46, *); C_δ, 29.4 (1.71, *); C_ϵ, 42.0 (2.99, 2.99)
F^{83}	121.8 (9.25)	176.1	57.5 (4.68)	39.7 (2.07, 2.84)	C_δ, * (7.27) C_ϵ, * (7.14); C_ζ, * (*)
R^{84}	124.3 (8.77)	176.1	56.6 (4.11)	29.9 (1.81, 1.91)	C_γ, 26.7 (1.40, *); C_δ, 43.3 (3.16, 3.16)
D^{85}	126.1 (8.73)	178.8	54.1 (5.14)	40.4 (3.03, 3.03)	
I^{86}	121.4 (8.96)	177.7	59.7 (4.64)	38.5 (2.61)	$C_{\gamma 1}$, 27.3 (1.45, 1.52); $C_{\gamma 2}$, 19.9 (0.79); $C_{\delta 1}$, 15.6 (0.83)
R^{87}	121.6 (8.77)	178.5	60.7 (3.64)	30.2 (1.72, 2.04)	C_γ, 28.6 (1.22, 1.33); C_δ, 43.3 (*, *)
G^{88}	103.3 (8.63)	174.9	45.8 (3.76, 3.40)		
F^{89}	115.7 (7.71)	176.4	56.4 (4.80)	39.7 (2.92, 3.29)	C_δ, * (7.23); C_ϵ, * (7.36); C_ζ, * (*)
I^{90}	122.2 (7.47)	176.0	63.8 (4.30)	37.4 (1.87)	$C_{\gamma 1}$, 28.6 (2.00, *); $C_{\gamma 2}$, 17.6 (0.73); $C_{\delta 1}$, 14.4 (0.88)
S^{91}	122.7 (8.48)	173.9	56.7 (4.88)	65.7 (3.35, 3.87)	
E^{92}	124.4 (8.98)	175.6	55.5 (4.60)	30.3 (1.88, 2.37)	C_γ, 36.7 (2.30, 2.34)
D^{93}	123.0 (8.46)	174.3	53.7 (4.86)	39.7 (2.50, 2.72)	
L^{94}	120.2 (7.72)	176.6	54.9 (3.81)	42.6 (0.82, 0.98)	C_γ, 26.3 (1.35); $C_{\delta 1}$, 25.5 (0.55); $C_{\delta 2}$, 25.5 (0.48)
E^{95}	120.1 (8.02)	178.7	58.7 (3.96)	29.1 (1.96, 2.04)	C_γ, 35.7 (2.21, 2.27)
G^{96}	114.0 (8.96)	174.0	46.1 (3.79, 4.07)		
V^{97}	120.1 (7.50)	175.2	62.16 (3.95)	33.3 (1.86)	$C_{\gamma 1}$, * (0.64); $C_{\gamma 2}$, 21.41 (0.79)
D^{98}	126.5 (8.27)	175.1	54.3 (4.41)	41.2 (2.48, 2.59)	
A^{99}	127.4 (8.19)	176.2	51.3 (4.90)	23.4 (1.30)	
T^{100}	117.6 (9.37)	172.0	61.0 (5.00)	72.3 (3.65)	$C_{\gamma 2}$, 23.0 (0.92)
L^{101}	127.3 (9.47)	175.0	53.1 (5.13)	46.6 (0.74, 1.33)	C_γ, 27.1 (*); $C_{\delta 1}$, 24.5 (0.50); $C_{\delta 2}$, 23.5 (0.30)
V^{102}	127.1 (9.05)	*	61.5 (4.73)	32.7 (2.06)	$C_{\gamma 1}$, 20.4 (0.92); $C_{\gamma 2}$, 21.1 (0.92)
V^{103}	130.3 (7.91)	171.3	61.4 (*)	34.4 (1.98)	$C_{\gamma 1}$, * (*); $C_{\gamma 2}$, * (0.66)
H^{104}	127.1 (8.92)	*	55.1 (5.49)	32.7 (2.98, 3.23)	$C_{\delta 2}$, * (6.78); $C_{\epsilon 1}$, * (7.57)
S^{105}	* (*)	172.9	57.9 (4.15)	63.5 (3.51, 3.90)	
N^{106}	117.0 (9.28)	174.3	55.2 (4.02)	38.1 (2.75, 3.02)	$N_{\delta 2}$, * (*, *)
N^{107}	116.9 (8.40)	174.0	53.6 (4.54)	37.0 (2.46, 2.53)	$N_{\delta 2}$, 110.5 (6.62, 7.32)
F^{108}	122.0 (8.11)	175.0	56.3 (4.69)	39.0 (2.93, 3.07)	C_δ, * (7.26); C_ϵ, * (7.57); C_ζ, * (*)
T^{109}	108.4 (7.10)	*	60.5 (4.35)	69.9 (4.10)	$C_{\gamma 2}$, 22.0 (0.83)
N^{110}	* (*)	174.8	54.8 (4.36)	39.4 (2.73, 2.73)	$N_{\delta 2}$, * (*, *)
T^{111}	115.5 (8.76)	*	64.3 (4.40)	71.1 (3.93)	$C_{\gamma 2}$, 20.5 (1.13)
I^{112}	127.7 (8.02)	175.2	60.8 (4.49)	39.4 (1.67)	$C_{\gamma 1}$, 27.8 (0.76, 1.49); $C_{\gamma 2}$, 18.9 (0.65); $C_{\delta 1}$, 13.8 (0.76)
L^{113}	128.2 (9.13)	175.7	53.0 (4.59)	43.9 (0.56, 1.38)	C_γ, 26.3 (1.14); $C_{\delta 1}$, 24.8 (0.42); $C_{\delta 2}$, 22.9 (0.73)
E^{114}	123.4 (8.60)	176.7	55.6 (4.64)	28.8 (1.92, 1.93)	C_γ, 35.2 (2.19, 2.29)
V^{115}	115.4 (7.47)	175.8	61.3 (4.17)	32.4 (2.06)	$C_{\gamma 1}$, 21.4 (-0.06); $C_{\gamma 2}$, 17.8 (0.06)

Continued on Next Page...

Table D.3: Apo HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
G^{116}	110.0 (7.95)		44.7 (3.88, 4.48)		
P^{117}		178.3	62.6 (4.75)	32.4 (1.83, 2.20)	C_γ, 27.4 (2.06, 2.12); C_δ, 49.7 (3.78, 3.90)
V^{118}	115.4 (8.25)	174.7	58.6 (5.26)	34.2 (1.95)	$C_{\gamma1}$, 23.8 (0.87); $C_{\gamma2}$, 19.1 (0.66)
T^{119}	113.3 (8.29)	172.5	59.9 (4.63)	71.4 (3.92)	$C_{\gamma2}$, 19.7 (1.05)
M^{120}	126.0 (9.08)	*	55.3 (4.77)	30.9 (1.92, 2.20)	C_γ, 31.3 (2.36, 2.71)
A^{121}	132.2 (9.28)	178.2	52.5 (4.44)	20.1 (1.32)	
G^{122}	107.5 (8.34)	*	46.2 (3.65, 4.13)		
L^{123}	* (*)	*	56.2 (4.99)	42.8 (1.40, 1.68)	C_γ, 27.3 (1.30); $C_{\delta1}$, 25.2 (0.77); $C_{\delta2}$, 24.7 (0.73)
I^{124}	124.6 (8.35)	*	58.0 (3.56)	42.2 (1.10)	$C_{\gamma1}$, 24.5 (0.90, 1.00); $C_{\gamma2}$, * (0.66); $C_{\delta1}$, 13.5 (-0.05)
N^{125}	122.5 (8.25)	173.5	53.3 (4.01)	39.4 (2.70, 3.25)	
L^{126}	130.3 (7.92)	*	56.9 (4.18)	39.4 (0.90, 1.77)	C_γ, * (1.16); $C_{\delta1}$, * (0.80); $C_{\delta2}$, * (0.80)
S^{127}	* (*)	*	* (*)	62.5 (3.89, *)	
S^{128}	* (*)	173.8	59.8 (3.96)	62.6 (4.21, 4.21)	
T^{129}	118.3 (7.9)		59.9 (4.70)	70.5 (4.11)	$C_{\gamma2}$, 21.0 (1.15)
P^{130}		*	(4.45)	31.8 (1.82, 2.23)	C_γ, * (2.04, 2.11); C_δ, 49.8 (3.67, 3.83)
T^{131}	* (*)	174.2	61.8 (*)	73.8 (*)	$C_{\gamma2}$, 21.5 (1.34)
N^{132}	123.2 (8.16)	*	52.6 (5.28)	39.0 (2.49, 2.65)	
R^{133}	* (8.17)	*	56.1 (3.72)	(1.72, 1.45)	C_γ, * (*, *); C_δ, 43.6 (3.13, *)
M^{134}	* (9.23)	*	54.4 (*)	* (1.68, *)	C_γ, 31.2 (2.15, 2.33); C_ϵ, * (1.92)
I^{135}	* (*)	174.6	60.7 (4.66)	41.1 (1.69)	$C_{\gamma1}$, 27.7 (1.86, *); $C_{\gamma2}$, 18.8 (0.88); $C_{\delta1}$, 15.7 (0.81)
R^{136}	127.0 (9.45)	174.7	54.1 (5.33)	35.6 (1.23, 1.62)	C_γ, 27.1 (1.36, 1.42); C_δ, 44.0 (2.97, 2.97)
Y^{137}	120.1 (7.36)	176.0	53.0 (4.90)	38.6 (2.65, 3.20)	C_δ, * (6.62); C_ϵ, * (6.47)
D^{138}	125.2 (9.33)	174.2	54.2 (4.96)	39.9 (2.78, 2.78)	
Y^{139}	123.5 (7.79)	173.0	58.2 (4.37)	42.5 (2.61, 3.19)	C_δ, * (6.96); C_ϵ, * (6.64)
A^{140}	129.4 (7.61)	176.0	51.1 (4.15)	17.5 (1.09)	
T^{141}	116.2 (7.61)	*	62.5 (4.22)	70.1 (4.08)	$C_{\gamma2}$, 19.0 (1.09)
K^{142}	* (*)	176.1	55.5 (4.34)	33.9 (2.36, *)	C_γ, 24.6 (*, *); C_ϵ, 45.1 (3.17, 3.17)
T^{143}	116.8 (8.22)	175.2	61.9 (4.31)	70.2 (4.16)	$C_{\gamma2}$, 21.5 (1.20)
G^{144}	112.4 (8.43)	*	45.5 (3.92, 4.63)		
Q^{145}	* (*)	*	54.7 (4.38)	27.5 (2.08, *)	C_γ, 34.5 (2.37, *)
C^{146}	109.6 (8.51)	*	63.7 (3.96)	27.6 (3.09, 3.52)	
G^{147}	109.8 (8.47)	172.2	45.3 (3.65, 5.50)		
G^{148}	103.8 (8.32)	171.4	46.2 (3.55, 4.31)		
V^{149}	123.2 (9.27)	174.4	62.8 (4.33)	34.0 (2.00)	$C_{\gamma1}$, 22.2 (0.84); $C_{\gamma2}$, 23.3 (1.04)
L^{150}	132.4 (8.90)	176.5	54.0 (5.13)	43.0 (1.38, 1.79)	C_γ, 26.8 (0.91); $C_{\delta1}$, 23.6 (0.29); $C_{\delta2}$, 23.6 (0.29)
C^{151}	124.4 (10.03)	170.5	56.3 (5.09)	33.1 (2.86, 3.22)	
A^{152}	119.5 (8.97)	176.0	51.4 (4.43)	22.3 (1.43)	
T^{153}	115.0 (8.21)	*	66.1 (3.16)	68.4 (3.99)	$C_{\gamma2}$, 21.6 (1.12)
G^{154}	* (*)	172.2	46.2 (2.66, 3.71)		

Continued on Next Page...

Table D.3: Apo HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
K^{155}	118.5 (7.67)	177.4	55.2 (5.06)	39.0 (0.77, 0.97)	C_γ, * (1.24, *); C_δ, 30.1 (1.53, 1.86); C_ϵ, 41.4 (3.03, 3.03)
I^{156}	120.8 (8.35)	174.5	60.9 (3.97)	40.5 (1.36)	$C_{\gamma1}$, 27.6 (1.46, 1.63); $C_{\gamma2}$, 17.8 (0.74); $C_{\delta1}$, 16.0 (0.98)
F^{157}	122.6 (8.82)	177.6	56.8 (5.52)	42.3 (2.49, 3.28)	C_δ, * (7.07); $C_{\epsilon2}$, * (6.96); C_ζ, * (*)
G^{158}	103.1 (7.82)	169.9	47.6 (4.16, 4.73)		
I^{159}	117.6 (8.55)	*	57.0 (5.33)	41.3 (1.93)	$C_{\gamma1}$, * (1.32, 1.47); $C_{\gamma2}$, 17.8 (0.94); $C_{\delta1}$, 11.9 (0.87)
H^{160}	* (9.21)	*	* (4.18)	* (3.14, 3.45)	$C_{\delta2}$, * (6.60); $C_{\epsilon1}$, * (*)
V^{161}	118.6 (9.26)	175.4	62.8 (5.22)	* (2.61)	$C_{\gamma1}$, 19.8 (0.79); $C_{\gamma2}$, 21.7 (1.21)
G^{162}	106.5 (6.91)	169.9	45.6 (4.00, 4.21)		
G^{163}	105.0 (8.32)	172.7	46.4 (4.33, 4.58)		
N^{164}	115.5 (8.19)	176.4	51.7 (5.15)	39.6 (2.85, 3.55)	$N_{\delta2}$, * (*, *)
G^{165}	111.9 (8.94)	174.3	45.8 (3.63, 4.54)		
R^{166}	118.8 (8.02)	174.7	56.4 (4.40)	32.1 (1.66, 1.85)	C_γ, 27.3 (1.59, 1.78); C_δ, 43.1 (3.14, 3.14)
Q^{167}	117.7 (8.46)	*	54.8 (4.97)	32.1 (1.99, *)	C_γ, 34.3 (2.12, 2.37); $N_{\epsilon2}$, * (*, *)
G^{168}	109.6 (8.56)	169.2	44.0 (3.42, 4.73)		
F^{169}	113.9 (8.01)	175.4	55.2 (5.72)	42.9 (2.22, 2.96)	C_δ, * (6.74); C_ϵ, * (6.90); C_ζ, * (7.02)
S^{170}	118.7 (9.78)	173.9	58.3 (5.77)	68.3 (3.50, 3.85)	
A^{171}	126.3 (9.30)	*	51.4 (3.98)	21.5 (1.34)	
Q^{172}	* (*)	174.9	55.4 (4.47)	32.0 (2.33, *)	C_γ, 33.9 (2.60, *); $N_{\epsilon2}$, * (*, *)
L^{173}	131.5 (7.98)	172.9	54.1 (4.61)	44.8 (1.25, 1.61)	C_γ, 25.0 (1.15); $C_{\delta1}$, 25.9 (0.79); $C_{\delta2}$, 24.2 (0.72)
K^{174}	115.0 (5.73)	176.9	54.0 (4.48)	34.6 (-0.30, 1.49)	C_γ, 24.9 (0.94, 1.03); C_δ, 29.0 (1.31, 1.50); C_ϵ, 41.4 (2.75, 2.75)
K^{175}	124.0 (9.20)	179.4	60.8 (3.82)	32.7 (1.76, 1.80)	C_γ, 26.2 (1.46, 1.46); C_δ, 29.9 (1.70, 1.70); C_ϵ, 42.0 (3.05, 3.05)
Q^{176}	114.3 (8.63)	176.4	57.9 (4.02)	28.6 (1.97, 2.05)	C_γ, 33.8 (2.21, 2.21); C_δ, 180.1 $N_{\epsilon2}$, 111.4 (7.05, 7.56)
Y^{177}	116.1 (6.90)	176.5	55.4 (4.46)	35.3 (1.95, 2.29)	C_δ, * (6.75); C_ϵ, * (7.02)
F^{178}	114.1 (7.60)	174.6	56.1 (5.10)	40.6 (2.55, 3.46)	C_δ, * (7.33); C_ϵ, * (7.19); C_ζ, * (7.02)
V^{179}	117.7 (7.28)	175.9	61.8 (4.13)	33.0 (2.05)	$C_{\gamma1}$, 20.2 (0.92); $C_{\gamma2}$, 21.2 (0.92)
E^{180}	125.4 (8.63)	176.3	56.6 (4.28)	30.3 (1.94, 2.04)	C_γ, 36.6 (2.30, 2.30)
K^{181}	123.7 (8.47)	175.5	56.5 (4.29)	33.0 (1.72, 1.82)	C_γ, 24.5 (1.43, 1.43); C_δ, 29.0 (1.65, 1.65); C_ϵ, 42.0 (2.94, 2.94)
Q^{182}	127.5 (8.05)		57.4 (4.17)	30.6 (1.89, 2.07)	C_γ, 34.3 (2.26, *); C_δ, 181.2 $N_{\epsilon2}$, 111.9 (6.78, 7.53)

D.4 Inhibited HRV14-3C Chemical Shifts

Table D.4: **Chemical Shift Assignments for the Inhibited HRV14-3C Protease**[2]

Residue	N	C′	C_α	C_β	Other
P^{2}		177.3	64.7 (4.39)	32.4 (2.43, 2.43)	C_γ, 27.2 (1.93, 1.99); C_δ, 49.9 (3.68, 3.68)
N^{3}	115.0 (8.90)	176.4	53.2 (5.01)	38.8 (3.04, 3.04)	C_γ, *; $N_{\delta 2}$, 116.1 (7.17, 8.49)
T^{4}	117.1 (7.64)	174.7	67.7 (3.69)	68.7 (4.10)	$C_{\gamma 2}$, 22.0 (1.24)
E^{5}	120.7 (8.64)	179.2	60.3 (3.93)	29.0 (2.05, 2.05)	C_γ, 36.4 (2.25, 2.33)
F^{6}	121.3 (8.58)	177.4	61.3 (4.22)	39.5 (2.81, 3.29)	C_δ, 132.7 (7.08); C_ϵ, 131.4 (6.54); C_ζ, 128.9 (6.22)
A^{7}	122.5 (8.32)	179.1	55.7 (3.70)	19.0 (1.36)	
L^{8}	117.1 (8.59)	179.6	57.8 (3.91)	41.8 (1.32, 1.8)	C_γ, 26.7 (1.68); $C_{\delta 1}$, 25.6 (0.81); $C_{\delta 2}$, 22.8 (0.80)
S^{9}	115.7 (7.74)	177.0	61.3 (4.12)	62.4 (3.87, 3.88)	
L^{10}	121.1 (7.63)	176.3	58.0 (3.78)	41.8 (1.13, 1.74)	C_γ, 26.0 (0.85); $C_{\delta 1}$, 23.2 (0.46); $C_{\delta 2}$, 26.0 (0.59)
L^{11}	120.3 (8.13)	177.6	58.0 (3.84)	41.7 (1.64, *)	C_γ, 27.2 (1.35); $C_{\delta 1}$, 23.4 (0.62); $C_{\delta 2}$, 24.5 (0.72)
R^{12}	114.7 (7.75)	178.5	58.5 (4.13)	31.1 (1.86, 1.94)	C_γ, 27.8 (1.67, 1.75); C_δ, 43.2 (3.17, 3,17)
K^{13}	113.9 (7.66)	178.0	57.3 (4.69)	35.6 (1.82, 1.86)	C_γ, 25.8 (1.42, 1.47); C_δ, 28.9 (1.30, 1.47); C_ϵ, 41.98 (2.61, 2.97)
N^{14}	116.7 (8.67)	172.7	56.1 (5.10)	43.5 (2.08, 3.22)	C_γ 177.6; $N_{\delta 2}$, 113.0 (7.60, 8.03)
I^{15}	121.1 (7.34)	174.4	60.8 (5.10)	39.0 (1.86)	$C_{\gamma 1}$, 29.2 (0.98, 1.37); $C_{\gamma 2}$, 18.8 (1.05); $C_{\delta 1}$, 14.3 (0.69)
M^{16}	124.4 (8.73)	174.2	53.7 (5.04)	38.1 (1.88, 2.15)	C_γ, 30.9 (2.29, 2.43); C_ϵ, 18.8 (1.69)
T^{17}	118.5 (8.89)	174.9	62.9 (5.08)	69.6 (3.90)	$C_{\gamma 2}$, 23.2 (1.09)
I^{18}	130.6 (9.68)	174.9	58.7 (5.53)	39.5 (1.76)	$C_{\gamma 1}$, 27.8 (1.20, 1.48); $C_{\gamma 2}$, 16.5 (0.78); $C_{\delta 1}$, 12.8 (0.82)
T^{19}	124.3 (9.41)	175.9	61.8 (5.47)	69.5 (3.92)	$C_{\gamma 2}$, 21.7 (1.35)
T^{20}	119.6 (9.33)	*	59.2 (5.35)	71.3 (4.75)	$C_{\gamma 2}$, 21.6 (1.00)
S^{21}	* (*)	175.4	61.6 (4.37)	62.3 (4.07, 4.07)	
K^{22}	118.3 (8.34)	175.9	54.5 (4.56)	32.5 (1.84, 2.27)	C_γ, 24.6 (1.45, 1.62); C_δ, 27.7 (1.68, 1.75); C_ϵ, 42.21 (2.99, 2.99); N_ζ, * (7.07)
G^{23}	107.4 (7.61)	171.0	44.2 (3.80, 4.43)		
E^{24}	117.5 (8.26)	176.0	55.9 (4.92)	31.6 (1.62, 1.82)	C_γ, 36.4 (2.23, 2.28)
F^{25}	119.6 (9.45)	175.6	56.8 (4.50)	42.8 (2.65, 2.98)	C_δ, 130.8 (7.24); C_ϵ, 130.8 (7.11); C_ζ, 132.0 (7.54)
T^{26}	120.1 (8.68)	172.3	65.0 (4.63)	69.9 (3.80)	$C_{\gamma 2}$, 22.8 (1.36)
G^{27}	111.4 (9.28)	171.2	44.2 (3.22, 5.01)		
L^{28}	121.9 (7.36)	175.4	52.9 (4.96)	43.7 (1.22, 1.74)	C_γ, 26.5 (1.40); $C_{\delta 1}$, 22.7 (0.80); $C_{\delta 2}$, 26.4 (0.66)
G^{29}	116.8 (9.23)	172.2	43.3 (1.15, 3.14)		
I^{30}	126.8 (8.80)	173.9	59.8 (3.88)	37.0 (1.56)	$C_{\gamma 1}$, 25.1 (0.76, 1.42); $C_{\gamma 2}$, 16.7 (0.15); $C_{\delta 1}$, 8.9 (0.20)
H^{31}	108.5 (7.30)	172.0	54.8 (4.44)	30.0 (3.05, 3.75)	$C_{\delta 2}$, 118.9 (6.46); $C_{\epsilon 1}$, 138.3 (7.79)
D^{32}	122.8 (10.06)	174.1	57.4 (4.02)	40.9 (2.64, 2.68)	
R^{33}	125.7 (8.30)	175.0	56.5 (4.72)	29.1 (2.05, 2.05)	C_γ, 29.0 (1.59, 1.59); C_δ, 43.8 (3.06, 3.10); N_ϵ, * (6.54); N_η, *(6.22)
V^{34}	121.9 (9.01)	175.4	62.7 (5.05)	33.1 (1.92)	$C_{\gamma 1}$, 22.0 (0.99); $C_{\gamma 2}$, 22.0 (0.99)
C^{35}	123.8 (9.63)	171.6	54.9 (5.53)	33.2 (2.52, 2.57)	
V^{36}	114.4 (8.39)	173.4	59.1 (5.61)	36.1 (2.09)	$C_{\gamma 1}$, 22.7 (1.12); $C_{\gamma 2}$, 21.4 (1.16)
I^{37}	116.4 (8.29)		58.1 (4.94)	43.1 (1.60)	$C_{\gamma 1}$, 26.3 (0.87, 1.31); $C_{\gamma 2}$, 17.6 (0.85); $C_{\delta 1}$, 14.8 (0.52)

Continued on Next Page...

Table D.4: Inhibited HRV14-3C chemical shifts – Continued

Residue	N ◦	C′	C_α	C_β	Other
P^{38}		177.4	62.2 (5.06)	32.5 (*, *)	C_γ, 29.5 (*, *); C_δ, * (3.76, 3.76)
T^{39}	124.1 (7.56)	*	65.0 (4.43)	69.0 (3.95)	$C_{\gamma 2}$, 20.0 (1.02)
H^{40}	* (7.93)	174.5	57.6 (4.28)	28.9 (2.77, 2.77)	$C_{\delta 2}$, 118.4 (6.62); $C_\epsilon 1$, 135.7 (7.82)
A^{41}	120.1 (7.31)	177.0	54.2 (3.50)	17.8 (0.18)	
Q^{42}	110.9 (8.00)		55.8 (3.95)	28.0 (1.80, 1.80)	C_γ, 34.0 (2.16, 2.16); C_δ, 180.2; $N_{\epsilon 2}$, 111.3 (6.78, 7.32)
P^{43}		176.8	62.9 (4.31)	32.0 (1.67, 1.82)	C_γ, 28.1 (2.05, 2.05); C_δ, 50.0 (3.73, 3.73)
G^{44}	112.8 (9.13)	173.1	44.0 (3.74, 4.38)		
D^{45}	121.2 (8.41)	175.4	56.1 (4.54)	41.3 (2.48, 2.86)	
D^{46}	114.7 (7.78)	173.8	52.6 (5.22)	44.7 (2.32, 2.58)	
V^{47}	111.5 (8.77)	174.7	59.1 (4.70)	36.2 (2.10)	$C_{\gamma 1}$, 22.3 (0.72); $C_{\gamma 2}$, 18.0 (0.50)
L^{48}	116.1 (7.96)	177.2	53.4 (5.28)	42.5 (0.88, 1.91)	C_γ, 26.9 (1.56); $C_{\delta 1}$, 22.3 (0.68); $C_{\delta 2}$, 25.0 (0.84)
V^{49}	122.8 (9.38)	176.6	61.4 (4.57)	32.6 (2.07)	$C_{\gamma 1}$, 22.0 (0.85); $C_{\gamma 2}$, 22.0 (0.85)
N^{50}	128.9 (9.87)	175.7	54.3 (4.47)	37.2 (3.09, 3.18)	$N_{\delta 2}$, 113.5 (7.29, 7.72)
G^{51}	101.2 (8.27)	173.1	45.8 (3.52, 4.31)		
Q^{52}	120.7 (7.77)	175.2	54.0 (4.60)	30.7 (2.03, 2.15)	C_γ, 33.61 (2.30, 2.41)
K^{53}	127.8 (8.84)	175.5	58.2 (4.45)	32.2 (1.77, 1.77)	C_γ, 24.6 (1.20, 1.20); C_δ, 29.0 (1.55, 1.55); C_ϵ, 41.8 (2.81, 2.81)
I^{54}	131.0 (9.35)	174.4	58.3 (4.50)	40.3 (1.81)	$C_{\gamma 1}$, 27.2 (1.31, 1.58); $C_{\gamma 2}$, 18.8 (0.97); $C_{\delta 1}$, 10.9 (0.81)
R^{55}	126.3 (8.87)	175.0	56.9 (4.53)	31.4 (1.90, 1.96)	C_γ, 27.6 (1.60, 1.80); C_δ, 43.2 (3.27, 3.36)
V^{56}	125.2 (8.45)	175.9	62.2 (4.04)	32.2 (2.04)	$C_{\gamma 1}$, 22.9 (0.80); $C_{\gamma 2}$, 23.9 (0.84)
K^{57}	129.1 (8.85)	175.9	57.9 (4.17)	33.5 (1.46, 1.70)	C_γ, 25.0 (1.38, 1.40); C_δ, 28.9 (1.64, 1.64); C_ϵ, 41.4 (2.88, 2.94)
D^{58}	114.4 (7.61)	173.0	53.6 (4.80)	44.6 (2.54, 2.66)	
K^{59}	118.1 (8.16)	174.4	55.0 (5.35)	35.9 (1.40, 1.71)	C_γ, 23.2 (1.24, 1.30); C_δ, 29.7 (1.31, 1.47); C_ϵ, 41.6 (2.89, 3.05)
Y^{60}	119.5 (8.39)	173.0	56.4 (4.82)	40.3 (2.90, 3.12)	C_δ, 133.3 (6.92); C_ϵ, 118.0 (6.74)
K^{61}	123.7 (8.52)	176.0	55.2 (4.71)	33.6 (1.48, 1.71)	C_γ, 24.9 (1.34, 1.34); C_δ, 29.5 (1.60, 1.67); C_ϵ, 41.8 (2.89, 2.89)
L^{62}	123.2 (8.70)	176.8	55.5 (4.54)	42.2 (0.99, 1.62)	C_γ, 27.2 (1.33); $C_{\delta 1}$, 24.6 (0.11); $C_{\delta 2}$, 23.0 (0.52)
V^{63}	117.5 (8.10)	174.3	59.3 (4.68)	36.1 (1.80)	$C_{\gamma 1}$, 18.4 (0.95); $C_{\gamma 2}$, 22.5 (0.85)
D^{64}	125.8 (8.91)		53.2 (4.64)	41.3 (2.56, 3.35)	
P^{65}		177.3	65.5 (4.43)	32.0 (2.48, 2.50)	C_γ, 28.0 (1.87, 2.04); C_δ, 51.8 (3.62, 3.72)
E^{66}	117.0 (8.20)	176.0	55.7 (4.39)	29.4 (2.30, 2.30)	C_γ, 37.4 (2.18, 2.18)
N^{67}	114.8 (8.68)	174.6	54.4 (4.18)	37.5 (2.85, 3.16)	C_γ, 178.5; $N_{\delta 2}$, 112.2 (6.74, 7.49)
I^{68}	121.7 (8.48)	176.1	60.8 (4.03)	36.1 (1.95)	$C_{\gamma 1}$, 27.9 (1.02, 1.47); $C_{\gamma 2}$, 18.2 (0.88); $C_{\delta 1}$, 11.8 (0.75)
N^{69}	124.2 (8.49)	174.4	56.1 (4.33)	40.6 (2.92, 3.39)	
L^{70}	121.9 (7.15)	176.4	54.3 (4.84)	41.9 (1.38, 1.75)	C_γ, 26.7 (*); $C_{\delta 1}$, 25.5 (0.65); $C_{\delta 2}$, 23.5 (0.93)
E^{71}	116.5 (7.87)	174.4	55.1 (4.11)	29.8 (2.45, 2.45)	C_γ, 38.7 (1.81, 3.43)
L^{72}	116.5 (6.49)	175.0	53.1 (4.78)	47.5 (1.00, 1.41)	C_γ, 27.3 (0.95); $C_{\delta 1}$, 24.3 (0.83); $C_{\delta 2}$, 24.3 (0.83)
T^{73}	118.9 (8.60)	171.4	62.0 (4.93)	71.7 (3.63)	$C_{\gamma 2}$, 20.8 (0.84)
V^{74}	127.3 (8.68)	175.7	61.6 (4.47)	32.4 (1.44)	$C_{\gamma 1}$, 22.7 (1.03); $C_{\gamma 2}$, 21.2 (0.79)
L^{75}	127.0 (9.34)	175.8	53.4 (5.07)	44.8 (1.30, 2.00)	C_γ, 26.1 (1.61); $C_{\delta 1}$, 25.1 (0.67); $C_{\delta 2}$, 24.9 (0.73)
T^{76}	117.5 (8.62)	174.9	62.0 (5.02)	68.9 (3.98)	$C_{\gamma 2}$, 21.8 (1.18)

Continued on Next Page...

Table D.4: Inhibited HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
L^{77}	126.4 (9.24)	176.2	53.9 (4.59)	44.2 (1.28, 1.77)	C_γ, 26.9 (1.86); $C_{\delta1}$, 22.0 (0.65); $C_{\delta2}$, 26.1 (0.93)
D^{78}	121.3 (8.54)	173.5	52.5 (4.78)	40.3 (2.09, 3.18)	
R^{79}	119.1 (7.25)	*	54.2 (4.72)	33.6 (1.79, 1.86)	C_γ, 26.5 (1.45, 1.52); C_δ, 44.3 (2.87, 3.05)
N^{80}	* (*)	176.3	55.6 (4.47)	39.0 (2.70, 2.70)	C_γ, 176.4; $N_{\delta2}$, 113.5 (6.92, 7.61)
E^{81}	120.2 (8.28)	175.2	57.7 (3.99)	30.2 (1.69, 1.71)	C_γ, 36.5 (2.16, 2.16)
K^{82}	120.7 (8.34)	177.4	54.9 (4.21)	35.1 (1.60, 1.74)	C_γ, 25.7 (1.42, 1.47); C_δ, 29.4 (1.76, 1.76); C_ϵ, 41.7 (3.00, 3.00)
F^{83}	121.9 (9.24)	176.1	57.3 (4.69)	39.9 (2.07, 2.88)	C_δ, 132.7 (7.22); C_ϵ, * (7.08); C_ζ, * (*)
R^{84}	124.3 (8.79)	176.0	56.6 (4.14)	29.7 (1.79, 1.83)	C_γ, 26.7 (1.43, 1.46); C_δ, 43.2 (3.13, 3.13)
D^{85}	126.1 (8.68)	178.8	54.0 (5.13)	40.4 (3.03, 3.03)	
I^{86}	121.3 (8.94)	177.7	59.8 (4.62)	38.3 (2.63)	$C_{\gamma1}$, 27.2 (1.47, 1.50); $C_{\gamma2}$, 20.0 (0.80); $C_{\delta1}$, 15.2 (0.80)
R^{87}	121.6 (8.76)	178.5	60.8 (3.63)	30.04 (1.72, 2.03)	C_γ, 28.8 (*, *); C_δ, 43.5 (2.98, 2.98); N_ϵ, * (7.38)
G^{88}	103.2 (8.62)	174.9	45.9 (3.46, 3.79)		
F^{89}	115.6 (7.66)	176.4	56.3 (4.75)	39.6 (2.92, 3.27)	C_δ, 131.4 (7.24); C_ϵ, * (6.96)
I^{90}	122.1 (7.44)	175.9	63.7 (4.26)	37.0 (1.86)	$C_{\gamma1}$, 28.6 (0.74, 1.96); $C_{\gamma2}$, 17.6 (0.75); $C_{\delta1}$, 14.3 (0.85)
S^{91}	122.7 (8.43)	173.9	56.7 (4.88)	65.7 (3.34, 3.88)	
E^{92}	124.3 (8.95)	175.5	55.4 (4.56)	30.3 (2.34, 2.40)	C_γ, 36.3 (2.06, 2.35)
D^{93}	123.1 (8.42)	174.3	53.6 (4.82)	39.9 (2.48, 2.77)	
L^{94}	120.3 (7.69)	176.7	55.0 (3.77)	42.5 (0.87, 1.02)	C_γ, 26.3 (1.35); $C_{\delta1}$, 25.3 (0.59); $C_{\delta2}$, 25.3 (0.50)
E^{95}	119.3 (7.95)	178.8	58.7 (3.95)	29.3 (1.95, 2.01)	C_γ, 35.8 (2.28, 2.35)
G^{96}	113.4 (8.92)	174.0	46.1 (3.73, 4.06)		
V^{97}	119.7 (7.42)	175.2	62.0 (3.96)	33.2 (1.87)	$C_{\gamma1}$, 21.9 (0.80); $C_{\gamma2}$, 21.9 (0.80)
D^{98}	126.7 (8.28)	174.9	54.4 (4.34)	41.2 (2.52, 2.61)	
A^{99}	127.4 (8.07)	176.1	51.1 (4.84)	23.3 (1.28)	
T^{100}	118.5 (9.38)	171.8	61.1 (4.96)	72.3 (3.63)	$C_{\gamma2}$, 23.23 (0.95)
L^{101}	127.6 (9.46)	174.9	53.0 (5.14)	46.8 (0.72, 1.39)	C_γ, 27.3 (0.95); $C_{\delta1}$, 24.3 (0.60); $C_{\delta2}$, 25.1 (0.37)
V^{102}	127.6 (8.96)	174.2	61.7 (4.67)	32.7 (1.98)	$C_{\gamma1}$, 22.8 (0.80); $C_{\gamma2}$, 22.8 (0.85)
V^{103}	128.7 (9.54)	171.8	61.5 (4.14)	34.6 (1.93)	$C_{\gamma1}$, 24.8 (1.26); $C_{\gamma2}$, 21.7 (0.66)
H^{104}	127.5 (8.98)	172.0	55.0 (5.46)	32.8 (2.45, 2.97)	$C_{\delta2}$, 116.5 (6.72); $C_{\epsilon1}$, 138.3 (7.57)
S^{105}	111.2 (7.30)	172.8	58.0 (4.11)	63.2 (3.50, 3.98)	
N^{106}	117.7 (9.41)	174.0	55.5 (4.02)	38.1 (2.79, 3.04)	
N^{107}	117.1 (8.41)	173.6	53.7 (4.54)	36.3 (2.36, 2.45)	C_γ, 175.7; $N_{\delta2}$, 109.8 (6.51, 7.40)
F^{108}	122.2 (8.08)	175.0	56.1 (4.75)	38.4 (2.95, 3.15)	C_δ, 132.3 (7.19); C_ϵ, 132.0 (7.53) C_ζ, 128.9 (6.94)
T^{109}	108.0 (7.05)	*	60.4 (4.33)	* (4.10)	$C_{\gamma2}$, 22.0 (0.82)
N^{110}	* (*)	174.7	54.9 (4.29)	39.1 (2.74, 2.74)	C_γ, *; $N_{\delta2}$, 114.4 (6.88, 7.66)
T^{111}	115.4 (8.84)	171.5	64.7 (4.38)	71.3 (3.93)	$C_{\gamma2}$, 20.4 (1.11)
I^{112}	127.8 (8.50)	175.3	60.7 (4.55)	39.1 (1.74)	$C_{\gamma1}$, 27.8 (1.50, 1.77); $C_{\gamma2}$, 18.8 (0.65); $C_{\delta1}$, 13.3 (0.77)
L^{113}	127.8 (9.10)	175.8	52.7 (4.59)	43.6 (0.65, 1.35)	C_γ, 24.6 (0.70); $C_{\delta1}$, 24.9 (0.45); $C_{\delta2}$, 22.3 (0.75)
E^{114}	123.1 (8.52)	176.5	55.1 (4.72)	27.8 (1.90, 2.08)	C_γ, 34.9 (2.18, 2.29)
V^{115}	114.5 (7.3)	175.8	61.2 (4.19)	32.3 (2.04)	$C_{\gamma1}$, 21.4 (-0.03); $C_{\gamma2}$, 17.8 (0.06)

Continued on Next Page...

Table D.4: Inhibited HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
G^{116}	109.8 (7.84)		45.2 (3.82, 4.43)		
P^{117}		178.3	62.5 (4.81)	32.4 (1.85, 2.19)	C_γ, 27.5 (2.05, 2.11); C_δ, 49.9 (3.67, 3.84)
V^{118}	115.1 (8.25)	174.7	58.7 (5.32)	34.4 (1.93)	$C_{\gamma 1}$, 23.9 (0.88); $C_{\gamma 2}$, 19.1 (0.62)
T^{119}	113.3 (8.27)	172.5	59.8 (4.59)	71.1 (3.85)	$C_{\gamma 2}$, 19.7 (1.02)
M^{120}	126.14 (8.99)	175.1	55.4 (4.71)	30.1 (1.95, 2.20)	C_γ, 31.4 (2.34, 2.71)
A^{121}	132.6 (9.37)	178.1	52.5 (4.50)	20.0 (1.34)	
G^{122}	107.2 (8.23)	171.5	46.4 (3.56, 4.21)		
L^{123}	128.3 (8.01)	177.2	55.2 (4.92)	43.7 (1.36, 1.60)	C_γ, 26.8 (1.32); $C_{\delta 1}$, 24.9 (0.74); $C_{\delta 2}$, 25.5 (0.78)
I^{124}	123.1 (8.86)	173.8	57.9 (4.50)	41.2 (1.35)	$C_{\gamma 1}$, 25.1 (0.66, 0.73); $C_{\gamma 2}$, 16.0 (0.18); $C_{\delta 1}$, 12.6 (-0.06)
N^{125}	120.6 (8.19)	173.4	52.5 (4.61)	38.9 (2.47, 2.58)	
L^{126}	128.2 (7.94)	174.9	52.7 (4.40)	42.3 (1.12, 1.64)	C_γ, 27.1 (1.12); $C_{\delta 1}$, 24.8 (0.58); $C_{\delta 2}$, 26.3 (0.75)
S^{127}	124.9 (9.51)	175.7	60.7 (3.83)	62.7 (3.82, 3.91)	
S^{128}	107.9 (8.43)	173.4	60.7 (4.02)	62.4 (4.15, 4.12)	
T^{129}	118.9 (7.97)		60.0 (4.71)	70.5 (4.32)	$C_{\gamma 2}$, 21.0 (1.23)
P^{130}		175.8	64.1 (4.41)	31.6 (1.94, 2.40)	C_γ, 27.3 (2.05, 2.05); C_δ, 49.7 (3.88, 3.88)
T^{131}	123.9 (9.19)	171.8	62.5 (5.23)	73.1 (3.80)	$C_{\gamma 2}$, 22.6 (1.21)
N^{132}	123.9 (8.20)	173.8	49.8 (5.40)	41.9 (2.42, 2.72)	C_γ, 177.0; $N_{\delta 2}$, 112.4 (6.90, 7.56)
R^{133}	111.6 (8.11)	173.2	57.4 (3.78)	27.0 (2.24, 2.24)	C_γ, 28.4 (*, *); C_δ, 44.2 (3.26, 3.24)
M^{134}	113.3 (8.41)	176.6	54.7 (5.15)	38.1 (1.57, 1.87)	C_γ, 32.1 (2.24, 2.34); C_ϵ, 20.1 (1.94)
I^{135}	126.1 (9.90)	174.5	60.9 (4.59)	41.0 (1.67)	$C_{\gamma 1}$, 27.7 (0.95, 1.85); $C_{\gamma 2}$, 18.6 (0.89); $C_{\delta 1}$, 15.7 (0.83)
R^{136}	127.7 (9.47)	174.5	54.1 (5.09)	35.7 (1.30, 1.60)	C_γ, 26.5 (0.77, 1.22); C_δ, 44.0 (2.86, 2.86)
Y^{137}	120.7 (7.45)	176.3	52.6 (4.89)	38.3 (2.57, 3.14)	C_δ, 133.3 (6.64); C_ϵ, 117.7 (6.63)
D^{138}	126.8 (9.59)	173.6	53.7 (4.99)	39.6 (2.75, 2.75)	
Y^{139}	122.8 (7.62)	173.4	57.8 (4.38)	43.5 (2.65, 3.18)	C_δ, 133.6 (6.98); $C_{\epsilon 1}$, 117.8 (6.68)
A^{140}	129.4 (7.39)	175.7	50.6 (4.03)	16.8 (1.03)	
T^{141}	117.1 (7.87)	175.7	63.0 (4.24)	69.4 (3.94)	$C_{\gamma 2}$, 22.2 (1.13)
K^{142}	120.1 (7.75)	176.2	54.6 (4.70)	37.8 (1.20, 2.90)	C_γ, 24.6 (1.32, 1.44); C_δ, 29.1 (1.69, 1.74); C_ϵ, 42.3 (3.04, 3.04)
T^{143}	113.3 (8.30)	174.6	65.3 (3.80)	69.4 (4.21)	$C_{\gamma 2}$, 21.9 (1.41)
G^{144}	109.7 (8.07)	176.0	45.2 (4.03, 4.63)		
Q^{145}	115.1 (7.29)	176.4	54.6 (4.33)	27.8 (2.12, 1.85)	C_γ, 34.3 (2.30, 2.36); C_δ, *; $N_{\epsilon 2}$, 109.9 (6.90, 8.17)
C^{146}	114.1 (7.85)	175.7	63.3 (4.66)	36.2 (3.30, 3.54)	
G^{147}	115.0 (10.64)	172.1	45.1 (3.60, 5.82)		
G^{148}	103.2 (8.06)	171.8	46.4 (3.35, 4.37)		
V^{149}	124.4 (9.33)	174.5	63.1 (4.31)	34.1 (1.98)	$C_{\gamma 1}$, 22.5 (0.88); $C_{\gamma 2}$, 23.8 (1.10)
L^{150}	132.9 (8.95)	176.4	54.1 (5.08)	42.9 (1.42, 1.80)	C_γ, 26.6 (0.95); $C_{\delta 1}$, 24.0 (0.24); $C_{\delta 2}$, 23.7 (0.24)
C^{151}	124.3 (10.05)	170.6	56.4 (5.11)	33.3 (2.92, 3.21)	
A^{152}	119.3 (8.98)	175.9	51.4 (4.39)	22.4 (1.44)	
T^{153}	114.8 (8.12)	*	66.2 (3.15)	68.6 (3.98)	$C_{\gamma 2}$, 21.4 (1.110)
G^{154}	* (*)	172.1	46.1 (2.67, 3.73)		

Continued on Next Page...

Table D.4: Inhibited HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
K^{155}	118.7 (7.67)	177.5	55.2 (5.07)	39.2 (0.80, 0.98)	C_γ, 25.5 (0.75, 1.22); C_δ, 30.4 (1.51, 1.88); C_ϵ, 42.6 (3.04. 3.04)
I^{156}	121.1 (8.30)	174.5	60.9 (3.97)	40.7 (1.33)	$C_{\gamma 1}$, 27.5 (1.46, 1.49); $C_{\gamma 2}$, 17.8 (0.79); $C_{\delta 1}$, 15.7 (0.96)
F^{157}	122.2 (8.84)	177.7	56.9 (5.53)	42.3 (2.49, 3.30)	C_δ, 130.8 (7.13); C_ϵ, 130.8 (6.98); C_ζ, * (7.60)
G^{158}	103.5 (7.79)	170.3	47.8 (4.21, 4.82)		
I^{159}	114.8 (8.5)	175.4	57.4 (5.64)	41.8 (1.90)	$C_{\gamma 1}$, 26.1 (1.23, 1.57); $C_{\gamma 2}$, 18.5 (1.00); $C_{\delta 1}$, 13.0 (0.88)
H^{160}	125.8 (10.47)	177.6	60.0 (4.75)	31.8 (3.18, 3.27)	$C_{\delta 2}$, 115.8 (6.73); $C_{\epsilon 1}$, 141.4 (8.40)
V^{161}	117.9 (9.29)	175.8	61.5 (5.20)	35.7 (2.72)	$C_{\gamma 1}$, 21.9 (1.30); $C_{\gamma 2}$, 20.1 (0.73)
G^{162}	105.8 (7.28)	170.4	46.6 (3.66, 4.94)		
G^{163}	103.6 (8.68)	172.9	46.9 (4.32, 4.87)		
N^{164}	117.4 (8.16)	175.8	51.9 (5.37)	40.1 (2.77, 3.56)	
G^{165}	111.6 (9.13)	174.3	45.5 (3.65, 4.66)		
R^{166}	117.5 (7.74)	174.1	56.5 (4.40)	32.8 (1.67, 1.83)	C_γ, 27.2 (1.54, 1.54); C_δ, 43.1 (3.14, 3.14)
Q^{167}	117.0 (8.43)	175.4	54.4 (4.97)	32.9 (1.97, 2.05)	C_γ, 34.6 (2.08, 2.44); C_δ, 180.4; $N_{\epsilon 2}$, 112.6 (6.63, 7.83)
G^{168}	109.8 (8.56)	169.5	44.6 (3.66, 5.10)		
F^{169}	113.1 (8.26)	175.4	55.7 (5.68)	43.4 (1.99, 2.95)	C_δ, 131.41 (6.55); C_ϵ, 130.2 (6.78); C_ζ, * (7.11)
S^{170}	119.8 (9.86)	173.9	58.6 (5.70)	68.5 (3.55, 3.92)	
A^{171}	125.5 (9.38)	177.4	51.7 (3.95)	20.9 (1.48)	
Q^{172}	124.2 (7.94)	176.7	57.2 (4.53)	30.2 (1.92, 2.66)	C_γ, 34.1 (2.10, 2.71); C_δ, 180.2; $N_{\epsilon 2}$, 112.3 (6.91, 7.59)
L^{173}	125.1 (8.51)	172.7	54.2 (4.59)	44.7 (1.24, 1.61)	C_γ, 24.9 (1.21); $C_{\delta 1}$, 27.5 (0.80); $C_{\delta 2}$, 26.3 (0.79)
K^{174}	115.0 (5.69)	176.9	53.9 (4.49)	34.7 (-0.28, 1.50)	C_γ, 25.0 (0.94, 1.05); C_δ, 29.3 (1.30, 1.50); C_ϵ, 41.2 (2.77, 2.77)
K^{175}	123.9 (9.17)	179.5	60.8 (3.81)	32.7 (1.77, 1.80)	C_γ, 26.3 (1.44, 1.46); C_δ, 29.7 (1.70, 1.70); C_ϵ, 41.7 (3.04, 3.04)
Q^{176}	114.3 (8.60)	176.4	58.0 (4.04)	28.4 (1.94, 2.03)	C_γ, 33.5 (2.25, 2.25); C_δ, 180.1; $N_{\epsilon 2}$, 112.2 (7.03, 7.56)
Y^{177}	116.0 (6.87)	176.4	55.4 (4.44)	35.4 (1.96, 2.30)	C_δ, 130.2 (6.73); C_ϵ, 119.0 (7.08)
F^{178}	114.0 (7.51)	174.7	56.1 (5.08)	40.7 (2.65, 3.45)	C_δ, 133.3 (7.34) C_ϵ, 129.0 (7.05); C_ζ, 130.8 (7.15)
V^{179}	118.0 (7.31)	175.9	62.0 (4.11)	33.0 (2.07)	$C_{\gamma 1}$, 21.3 (*); $C_{\gamma 2}$, 20.4 (0.95)
E^{180}	125.1 (8.58)	176.2	56.8 (4.26)	30.6 (1.95, 2.09)	C_γ, 36.5 (2.30, 2.35)
K^{181}	123.6 (8.43)	175.5	56.5 (4.26)	33.2 (1.78, 1.89)	C_γ, 24.6 (1.45, 1.45); C_δ, 29.0 (1.70, 1.70); C_ϵ, 41.9 (2.59, 2.91)
Q^{182}	127.7 (8.04)		57.4 (4.16)	30.6 (1.91, 2.11)	C_γ, 34.3 (2.29, 2.30); C_δ, 181.2; $N_{\epsilon 2}$, 112.3 (6.75, 7.49)

Continued on Next Page...

Table D.4: Inhibited HRV14-3C chemical shifts – Continued

Residue	N	C′	C_α	C_β	Other
			HRV14-3C bound Acetyl-LEALFQ-Ethylpropionate inhibitor chemical shifts		
Ace					(3.35)
L^{P6}			(4.38)	(1.45, 1.62)	C_γ, * (1.31); $C_{\delta 1}$ * (0.88); $C_{\delta 2}$ * (0.90)
E^{P5}			(4.20)	(1.82, 1.88)	C_γ, * (2.20, 2.71)
A^{P4}			(4.32)	(0.85)	
L^{P3}	(8.33)		(4.90)	(1.51, 1.68)	C_γ, * (1.38); $C_{\delta 1}$, * (0.86); $C_{\delta 2}$, * (0.90)
F^{P2}	(6.78)		(4.97)	(2.43, 2.52)	C_δ, * (7.21); C_ϵ, * (7.08); C_ζ, * (7.27)
Q^{P1}	(7.57)		(4.61)	(1.90, 1.97)	C_γ, * (2.15, 2.24); N_ϵ, * (6.57, 6.90)
Eth^{H1}					(1.32)
Eth^{H2}					(3.95)
Eth^{H3}					(4.22)
Eth^{H4}					(3.62)
Eth^{H5}					(4.92)
$Eth^{H6,7,8}$					(1.22)

Appendix E

Protection Factor Calculations

Table E.1: Apo HRV14-3C K_{rc} Calculations

Res	L_{acid}	L_{base}	R_{acid}	R_{acid}	Antilog K_a[1]	Antilog K_b[2]	Antilog K_w[3]	K_{rc}[4] s^{-1}	$K_{rc}(T)$[5] s^{-1}
I^{15}	-0.91	-0.73	-0.13	0.32	4.79×10^{-7}	7.76	1.23×10^{-2}	7.77	2.73
M^{16}	-0.79	-0.07	-0.59	-0.23	2.19×10^{-7}	1.00×10^{1}	1.58×10^{-2}	1.00×10^{1}	3.09
T^{17}	-0.79	-0.07	-0.28	0.11	4.47×10^{-7}	2.19×10^{1}	3.47×10^{-2}	2.19×10^{1}	4.54
I^{18}	-0.91	-0.73	-0.44	-0.11	2.34×10^{-7}	2.88	4.57×10^{-3}	2.89	1.68
T^{19}	-0.79	-0.07	-0.59	-0.23	2.19×10^{-7}	1.00×10^{1}	1.58×10^{-2}	1.00×10^{1}	3.09
T^{20}	-0.79	-0.07	-0.44	-0.11	3.09×10^{-7}	1.32×10^{1}	2.09×10^{-2}	1.32×10^{1}	3.54
F^{25}	-0.52	0.24	0.31	-0.15	3.24×10^{-6}	8.13	1.29×10^{-2}	8.14	2.79
L^{28}	-0.57	0.58	0.22	0.17	2.34×10^{-6}	7.76	1.23×10^{-2}	7.77	2.73
G^{29}	-0.02	0.27	-0.13	-0.21	3.70×10^{-6}	2.29×10^{1}	3.63×10^{-2}	2.29×10^{1}	4.64
I^{30}	-0.91	-0.73	0.22	0.17	1.07×10^{-6}	5.50	8.71×10^{-3}	5.50	2.31
H^{31}	0	-0.1	-0.59	-0.23	1.85×10^{-6}	9.33	1.48×10^{-2}	9.35	2.99
C^{35}	-0.54	0.62	-0.3	-0.14	7.59×10^{-7}	6.03×10^{1}	9.55×10^{-2}	6.04×10^{1}	7.45
V^{47}	-0.74	-0.7	0.58	-0.18	3.63×10^{-6}	2.63	4.17×10^{-3}	2.63	1.61
L^{48}	-0.57	-0.58	-0.3	-0.14	7.08×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
V^{49}	-0.74	-0.7	-0.13	-0.21	7.08×10^{-7}	2.45	3.89×10^{-3}	2.46	1.55
Q^{52}	-0.47	0.06	0.22	0.17	2.95×10^{-6}	3.39×10^{1}	5.37×10^{-2}	3.39×10^{1}	5.62
I^{54}	-0.91	-0.73	-0.29	0.12	3.31×10^{-7}	4.90	7.76×10^{-3}	4.91 x 10	2.18
V^{56}	-0.74	-0.7	-0.32	0.22	4.57×10^{-7}	6.61	1.05×10^{-2}	6.62	2.52
Y^{60}	-0.41	-0.27	-0.29	0.12	1.05×10^{-6}	1.41×10^{1}	2.24×10^{-2}	1.41×10^{1}	3.66
T^{73}	-0.79	-0.07	-0.13	-0.21	6.31×10^{-7}	1.05×10^{1}	1.66×10^{-2}	1.05×10^{1}	3.16
V^{74}	-0.74	-0.7	-0.44	-0.11	3.47×10^{-7}	3.09	4.90×10^{-3}	3.10	1.74
T^{76}	-0.79	-0.07	-0.13	-0.21	6.31×10^{-7}	1.05×10^{1}	1.66×10^{-2}	1.05×10^{1}	3.16
L^{77}	-0.57	-0.58	-0.44	-0.11	5.13×10^{-7}	4.07	6.46×10^{-3}	4.08	1.99
D^{78}	0.9	-0.3	-0.13	-0.21	3.09×10^{-5}	6.17	9.77×10^{-3}	6.18	2.44
K^{82}	-0.56	-0.04	0.31	-0.15	2.95×10^{-6}	1.29×10^{1}	2.04×10^{-2}	1.29×10^{1}	3.50
F^{83}	-0.52	-0.24	-0.29	0.12	8.13×10^{-7}	1.51×10^{1}	2.40×10^{-2}	1.52×10^{1}	3.79
I^{86}	-0.91	-0.73	0.58	-0.18	2.45×10^{-6}	2.45	3.89×10^{-3}	2.46	1.55
R^{87}	-0.59	0.08	-0.59	-0.23	3.47×10^{-7}	1.41×10^{1}	2.24×10^{-2}	1.41×10^{1}	3.66
I^{90}	-0.91	-0.73	-0.43	0.06	2.40×10^{-7}	4.27	6.76×10^{-3}	4.27	2.04
S^{91}	-0.44	0.37	-0.59	-0.23	4.90×10^{-7}	2.75×10^{1}	4.37×10^{-2}	2.76×10^{1}	5.08
E^{95}	-0.9	-0.51	-0.13	-0.21	4.90×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
L^{101}	-0.57	-0.58	-0.44	-0.11	5.13×10^{-7}	4.07	6.46×10^{-3}	4.08	1.99
V^{102}	-0.74	-0.7	-0.13	-0.21	7.08×10^{-7}	2.45	3.89×10^{-3}	2.46	1.55

Continued on Next Page...

Table E.1: Apo HRV14-3C K_{rc} Calculations – Continued

Res	L_{acid}	L_{base}	R_{acid}	R_{acid}	Antilog K_a [1]	Antilog K_b [2]	Antilog K_w [3]	K_{rc} [4] s^{-1}	$K_{rc}(T)$ [5] s^{-1}
T[129]	-0.79	-0.07	-0.13	-0.21	6.31×10^{-7}	1.05×10^{1}	1.66×10^{-2}	1.05×10^{1}	3.16
Y[137]	-0.41	-0.27	-0.32	0.22	9.77×10^{-7}	1.78×10^{1}	2.82×10^{-2}	1.78×10^{1}	4.10
C[151]	-0.54	0.62	-0.13	-0.21	1.12×10^{-6}	5.13×10^{1}	8.13×10^{-2}	5.14×10^{1}	6.89
A[152]	0	0	-0.46	0.55	1.82×10^{-6}	7.08×10^{1}	1.12×10^{-1}	7.09×10^{1}	8.07
F[157]	-0.52	-0.24	-0.59	-0.23	4.07×10^{-7}	6.76	1.07×10^{-2}	6.77	2.55

Table E.2: **Inhibited HRV14-3C K_{rc} Calculations**

Res	L_{acid}	L_{base}	R_{acid}	R_{acid}	Antilog K_a [1]	Antilog K_b [2]	Antilog K_w [3]	K_{rc} [4] s^{-1}	$K_{rc}(T)$ [5] s^{-1}
F [6]	-0.52	-0.24	0.31	-0.15	3.24×10^{-6}	8.13	1.29×10^{-2}	8.14	2.79
S [9]	-0.44	0.37	-0.13	-0.21	1.41×10^{-6}	2.88×10^{1}	4.57×10^{-2}	2.89×10^{1}	5.20
L [11]	-0.57	-0.58	-0.13	-0.21	1.05×10^{-6}	3.24	5.13×10^{-3}	3.24	1.78
I [15]	-0.91	-0.73	-0.13	0.32	4.79×10^{-7}	7.76	1.23×10^{-2}	7.77	2.73
M [16]	-0.79	-0.07	-0.59	-0.23	2.19×10^{-7}	1.00×10^{1}	1.58×10^{-2}	1.00×10^{1}	3.09
I [18]	-0.91	-0.73	-0.44	-0.11	2.34×10^{-7}	2.88	4.57×10^{-3}	2.89	1.68
T [19]	-0.79	-0.07	-0.59	-0.23	2.19×10^{-7}	1.00×10^{1}	1.58×10^{-2}	1.00×10^{1}	3.09
F [25]	-0.52	-0.24	0.31	-0.15	3.24×10^{-6}	8.13	1.29×10^{-2}	8.14	2.79
T [26]	-0.79	-0.07	-0.43	0.06	3.16×10^{-7}	1.95×10^{1}	3.09×10^{-2}	1.95×10^{1}	4.29
G [27]	-0.02	0.27	-0.44	-0.11	1.81×10^{-6}	2.88×10^{1}	4.57×10^{-2}	2.89×10^{1}	5.20
L [28]	-0.57	-0.58	0.22	0.17	2.34×10^{-6}	7.76	1.23×10^{-2}	7.77	2.73
G [29]	-0.02	0.27	-0.13	-0.21	3.70×10^{-6}	2.29×10^{1}	3.63×10^{-2}	2.29×10^{1}	4.64
I [30]	-0.91	-0.73	0.22	0.17	1.07×10^{-6}	5.50	8.71×10^{-3}	5.50	2.31
H [31]	0	-0.1	-0.59	-0.23	1.35×10^{-6}	9.33	1.48×10^{-2}	9.35	2.99
V [34]	-0.74	-0.7	-0.32	0.22	4.57×10^{-7}	6.61	1.05×10^{-2}	6.62	2.52
C [35]	-0.54	0.62	-0.3	-0.14	7.59×10^{-7}	6.03×10^{1}	9.55×10^{-2}	6.04×10^{1}	7.45
V [36]	-0.74	-0.7	-0.46	0.55	3.31×10^{-7}	1.41×10^{1}	2.24×10^{-2}	1.41×10^{1}	3.66
I [37]	-0.91	-0.73	-0.3	-0.14	3.24×10^{-7}	2.69	4.27×10^{-3}	2.70	1.63
T [39]	-0.79	-0.07	-0.19	-0.24	5.50×10^{-7}	9.77	1.55×10^{-2}	9.79	3.06
A [41]	0	0	0	0.14	5.25×10^{-6}	2.75×10^{1}	4.37×10^{-2}	2.76×10^{1}	5.08
Q [42]	-0.47	0.06	0	0	1.78×10^{-6}	2.29×10^{1}	3.63×10^{-2}	2.29×10^{1}	4.64
V [47]	-0.74	-0.7	0.58	-0.18	3.63×10^{-6}	2.63	4.17×10^{-3}	2.63	1.61
L [48]	-0.57	-0.58	-0.3	-0.14	7.08×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
V [49]	-0.74	-0.7	-0.13	-0.21	7.08×10^{-7}	2.45	3.89×10^{-3}	2.46	1.55
Q [52]	-0.47	0.06	0.22	0.17	2.95×10^{-6}	3.39×10^{1}	5.37×10^{-2}	3.39×10^{1}	5.62
I [54]	-0.91	-0.73	-0.29	0.12	3.31×10^{-7}	4.90	7.76×10^{-3}	4.91	2.18
Y [60]	-0.41	-0.27	-0.29	0.12	1.05×10^{-6}	1.41×10^{1}	2.24×10^{-2}	1.41×10^{1}	3.66
L [62]	-0.57	-0.58	-0.29	0.12	7.24×10^{-7}	6.92	1.10×10^{-2}	6.93	2.58
E [71]	-0.9	-0.51	-0.13	-0.21	4.90×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
L [72]	-0.57	-0.58	0.31	-0.15	2.88×10^{-6}	3.72	5.89×10^{-3}	3.72	1.90
T [73]	-0.79	-0.07	-0.13	-0.21	6.31×10^{-7}	1.05×10^{1}	1.66×10^{-2}	1.05×10^{1}	3.16
V [74]	-0.74	-0.7	-0.44	-0.11	3.47×10^{-7}	3.09	4.90×10^{-3}	3.10	1.74
L [75]	-0.57	-0.58	-0.3	-0.14	7.08×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
T [76]	-0.79	-0.07	-0.13	-0.21	6.31×10^{-7}	1.05×10^{1}	1.66×10^{-2}	1.05×10^{1}	3.16
L [77]	-0.57	-0.58	-0.44	-0.11	5.13×10^{-7}	4.07	6.46×10^{-3}	4.08	1.99
I [78]	0.9	-0.3	-0.13	-0.21	3.09×10^{-5}	6.17	9.77×10^{-3}	6.18	2.44
K [82]	-0.56	-0.04	0.31	-0.15	2.95×10^{-6}	1.29×10^{1}	2.04×10^{-2}	1.29×10^{1}	3.50
F [83]	-0.52	-0.24	-0.29	0.12	8.13×10^{-7}	1.51×10^{1}	2.40×10^{-2}	1.52×10^{1}	3.79
I [86]	-0.91	-0.73	0.58	-0.18	2.45×10^{-6}	2.45	3.89×10^{-3}	2.46	1.55
R [87]	-0.59	0.08	-0.59	-0.23	3.47×10^{-7}	1.41×10^{1}	2.24×10^{-2}	1.41×10^{1}	3.66
I [90]	-0.91	-0.73	-0.43	0.06	2.40×10^{-7}	4.27	6.76×10^{-3}	4.27	2.04
S [91]	-0.44	0.37	-0.59	-0.23	4.90×10^{-7}	2.75×10^{1}	4.37×10^{-2}	2.76×10^{1}	5.08

Continued on Next Page...

Table E.2: Inhibited HRV14-3C K_{rc} Calculations – Continued

Res	L_{acid}	L_{base}	R_{acid}	R_{acid}	Antilog K_a[1]	Antilog K_b[2]	Antilog K_w[3]	K_{rc}[4] s^{-1}	$K_{rc}(T)$[5] s^{-1}
A 99	0	0	0.58	-0.18	2.00×10^{-5}	1.32×10^{1}	2.09×10^{-2}	1.32×10^{1}	3.54
T 100	-0.79	-0.07	0	0	8.51×10^{-7}	1.70×10^{1}	2.69×10^{-2}	1.70×10^{1}	4.01
L 101	-0.57	-0.58	-0.44	-0.11	5.13×10^{-7}	4.07	6.46×10^{-3}	4.08	1.99
V 102	-0.74	-0.7	-0.13	-0.21	7.08×10^{-7}	2.45	3.89×10^{-3}	2.46	1.55
V 103	-0.74	-0.7	-0.3	-0.14	4.79×10^{-7}	2.88	4.57×10^{-3}	2.89	1.68
L 113	-0.57	-0.58	-0.59	-0.23	3.63×10^{-7}	3.09	4.90×10^{-3}	3.10	1.74
V 115	-0.74	-0.7	0.31	-0.15	1.95×10^{-6}	2.82	4.47×10^{-3}	2.82	1.66
G 116	-0.02	0.27	-0.3	-0.14	2.50×10^{-6}	2.69×10^{1}	4.27×10^{-2}	2.70×10^{1}	5.02
V 118	-0.74	-0.7	-0.19	-0.24	6.17×10^{-7}	2.29	3.63×10^{-3}	2.29	1.50
A 121	0	0	-0.28	0.11	2.75×10^{-6}	2.57×10^{1}	4.07×10^{-2}	2.57×10^{1}	4.91
I 124	-0.91	-0.73	-0.13	-0.21	4.79×10^{-7}	2.29	3.63×10^{-3}	2.29	1.50
L 126	-0.57	-0.58	-0.13	0.32	1.05×10^{-6}	1.10×10^{1}	1.74×10^{-2}	1.10×10^{1}	3.23
T 131	-0.79	-0.07	-0.19	-0.24	5.50×10^{-7}	9.77	1.55×10^{-2}	9.79	3.06
N 132	-0.58	0.49	-0.44	-0.11	5.01×10^{-7}	4.7×109^{1}	7.59×10^{-2}	4.79×10^{1}	6.66
R 133	-0.59	0.08	-0.13	0.32	1.00×10^{-6}	5.01×10^{1}	7.94×10^{-2}	5.02×10^{1}	6.81
M 134	-0.64	-0.01	-0.32	0.22	5.75×10^{-7}	3.24×10^{1}	5.13×10^{-2}	3.24×10^{1}	5.50
I 135	-0.91	-0.73	-0.28	0.11	3.39×10^{-7}	4.79	7.59×10^{-3}	4.79	2.16
Y 137	-0.41	-0.27	-0.32	0.22	9.77×10^{-7}	1.78×10^{1}	2.82×10^{-2}	1.78×10^{1}	4.10
Y 139	-0.41	-0.27	0.58	-0.18	7.76×10^{-6}	7.08	1.12×10^{-2}	7.09	2.61
T 141	-0.79	-0.07	0	0	8.51×10^{-7}	1.70×10^{1}	2.69×10^{-2}	1.70×10^{1}	4.01
Q 145	-0.47	0.06	0.22	0.17	2.95×10^{-6}	3.39×10^{1}	5.37×10^{-2}	3.39×10^{1}	5.62
G 147	-0.02	0.27	-0.46	0.55	1.73×10^{-6}	1.32×10^{2}	2.09×10^{-1}	1.32×10^{2}	10.9
G 148	-0.02	0.27	0.22	0.17	8.28×10^{-6}	5.50×10^{1}	8.71×10^{-2}	5.50×10^{1}	7.13
V 149	-0.74	-0.7	0.22	0.17	1.58×10^{-6}	5.89	9.33×10^{-3}	5.90	2.39
L 150	-0.57	-0.58	-0.3	-0.14	7.08×10^{-7}	3.80	6.03×10^{-3}	3.81	1.93
C 151	-0.54	0.62	-0.13	-0.21	1.12×10^{-6}	5.13×10^{1}	8.13×10^{-2}	5.14×10^{1}	6.89
A 152	0	0	-0.46	0.55	1.82×10^{-6}	7.08×10^{1}	1.12×10^{-1}	7.09×10^{1}	8.07
K 155	-0.56	-0.04	0.22	0.17	2.40×10^{-6}	2.69×10^{1}	4.27×10^{-2}	2.70×10^{1}	5.02
F 157	-0.52	-0.24	-0.59	-0.23	4.07×10^{-7}	6.76	1.07×10^{-2}	6.77	2.55
G 158	-0.02	0.27	-0.43	0.06	1.85×10^{-6}	4.27×10^{1}	6.76×10^{-2}	4.27×10^{1}	6.29
I 159	-0.91	-0.73	0.22	0.17	1.07×10^{-6}	5.50	8.71×10^{-3}	5.50	2.31
H 160	0	-0.1	-0.59	-0.23	1.35×10^{-6}	9.33	1.48×10^{-2}	9.35	2.99
V 161	-0.74	-0.7	0	0.14	9.55×10^{-7}	5.50	8.71×10^{-3}	5.50	2.31
G 162	-0.02	0.27	-0.3	-0.14	2.50×10^{-6}	2.69×10^{1}	4.27×10^{-2}	2.70×10^{1}	5.02
G 163	-0.02	0.27	0.22	0.17	8.28×10^{-6}	5.50×10^{1}	8.71×10^{-2}	5.50×10^{1}	7.13
N 164	-0.58	0.49	0.22	0.17	2.29×10^{-6}	9.12×10^{1}	1.45×10^{-1}	9.13×10^{1}	9.13
Q 167	-0.47	0.06	-0.32	0.22	8.51×10^{-7}	3.80×10^{1}	6.03×10^{-2}	3.81×10^{1}	5.95
G 168	-0.02	0.27	-0.27	0.2	2.68×10^{-6}	5.89×10^{1}	9.33×10^{-2}	5.90×10^{1}	7.37
F 169	-0.52	-0.24	0.22	0.17	2.63×10^{-6}	1.70×10^{1}	2.69×10^{-2}	1.70×10^{1}	4.01
S 170	-0.44	0.37	-0.43	0.06	7.08×10^{-7}	5.37×10^{1}	8.51×10^{-2}	5.38×10^{1}	7.05
A 171	0	0	-0.39	0.3	2.14×10^{-6}	3.98×10^{1}	6.31×10^{-2}	3.99×10^{1}	6.08
Q 172	-0.47	0.06	-0.27	0.2	9.55×10^{-7}	3.63×10^{1}	5.75×10^{-2}	3.64×10^{1}	5.82
L 173	-0.57	-0.58	-0.27	0.2	7.59×10^{-7}	8.32	1.32×10^{-2}	8.33	2.83
K 174	-0.56	-0.04	-0.13	-0.21	1.07×10^{-6}	1.12×10^{1}	1.78×10^{-2}	1.12×10^{1}	3.27

[1] $\log K_a = \log K_{a,ref} + \log A_L + \log A_R - pD$; where $K_{a,ref}=1.62$ and pD=6.9

[2] $\log K_b = \log K_{b,ref} + \log B_L + \log B_R - \log[OD^-]$; where $K_{b,ref}=10.05$ and $\log[OD^-]$=6.9-15.65=-8.75

[3] $\log K_w = \log K_{w,ref} + \log B_L + \log B_R$; where $K_{w,ref}$=-1.5

[4] $K_{rc} = K_a + K_b + K_w = K_{a,ref}(A_L * A_R)[D^+] + K_{b,ref}(B_L * B_R)[OD^-] + K_{w,ref}(B_L * B_R)$

[5] $K_{rc}(T) = K_{rc}(293)\exp(-Ea[1/T - 1/293]/R)$; where Ea=17 kcal/mol and R=1.987 cal/K·mol

Table E.3: **Apo HRV14-3C K_{ex} and P_{factor} Calculations†.**

Res	Volumes at Timepoint t_{min}					$V = V_0^{-K_{ex}t} + Baseline$			K_{rc} s^{-1}	P_{factor}
	120[1]	181	256	568	6662	V_0	K_{ex} s^{-1}	Baseline		(K_{rc} / K_{ex})
I^{15}	0.16	0.27	0.22	0.18	0.04	2.46×10^{-1}	1.31×10^{-3}	4.11×10^{-2}	2.73	9.22×10^{5}
M^{16}	0.06	0.11	0.11	0.1	0.08	3.39×10^{-2}	4.90×10^{-4}	7.76×10^{-2}	3.09	2.41×10^{3}
I^{18}	0.08	0.16	0.16	0.16	0.13	4.83×10^{-2}	1.60×10^{-4}	1.14×10^{-1}	1.68	6.78×10^{4}
T^{19}	0.19	0.39	0.37	0.33	0.23	1.80×10^{-1}	1.18×10^{-3}	2.26×10^{-1}	3.09	8.29×10^{2}
T^{20}	0.35	0.69	0.62	0.6	0.39	2.93×10^{-1}	7.50×10^{-4}	3.90×10^{-1}	3.54	1.96×10^{3}
F^{25}	0.12	0.02	0.28	0.02	0.04	1.23×10^{-1}	1.94×10^{-3}	3.30×10^{-2}	2.79	2.16×10^{2}
L^{28}	0.1	0.17	0.16	0.14	0.00	1.93×10^{-1}	3.40×10^{-4}	2.00×10^{-2}	2.73	4.36×10^{3}
G^{29}	0.13	0.27	0.25	0.24	0.21	7.51×10^{-2}	1.96×10^{-3}	2.05×10^{-1}	4.64	8.72×10^{5}
I^{30}	0.33	0.6	0.58	0.61	0.55	6.00×10^{-1}	$<2.50 \times 10^{-6}$	0.00	2.31	$>1.00 \times 10^{6}$
H^{31}	0.17	0.33	0.3	0.26	0.16	1.82×10^{-1}	1.24×10^{-3}	1.58×10^{-1}	2.99	1.83×10^{3}
C^{35}	0.15	0.29	0.28	0.28	0.24	1.04×10^{-1}	1.10×10^{-4}	1.85×10^{-1}	7.45	7.42×10^{2}
V^{47}	0.05	0.06	0.03	0.03	0.01	4.69×10^{-2}	1.94×10^{-3}	1.13×10^{-2}	1.61	3.05×10^{3}
L^{48}	0.13	0.27	0.26	0.24	0.18	9.38×10^{-2}	9.80×10^{-4}	1.80×10^{-1}	1.93	1.05×10^{5}
V^{49}	0.29	0.59	0.57	0.55	0.55	7.95×10^{-2}	7.20×10^{-3}	5.50×10^{-1}	1.55	4.06×10^{3}
Q^{52}	0.29	0.5	0.38	0.32	0.06	4.64×10^{-1}	1.29×10^{-3}	5.94×10^{-2}	5.62	1.78×10^{2}
I^{54}	0.29	0.58	0.53	0.53	0.53	5.50×10^{-1}	$<2.50 \times 10^{-6}$	0.00	2.18	$>1.00 \times 10^{6}$
V^{56}	0.16	0.23	0.15	0.13	0.00	2.29×10^{-1}	1.38×10^{-3}	5.37×10^{-3}	2.52	1.83×10^{3}
T^{73}	0.22	0.42	0.38	0.05	0.02	6.82×10^{-1}	4.26×10^{-3}	7.00×10^{-4}	3.16	7.42×10^{2}
V^{74}	0.18	0.37	0.33	0.35	0.32	3.38×10^{-2}	5.70×10^{-4}	3.23×10^{-1}	1.74	3.05×10^{3}
T^{76}	0.27	0.54	0.53	0.54	0.47	3.54×10^{-1}	3.00×10^{-5}	1.86×10^{-1}	3.16	1.05×10^{5}
L^{77}	0.19	0.35	0.36	0.33	0.25	1.19×10^{-1}	4.90×10^{-4}	2.43×10^{-1}	1.99	4.06×10^{3}
D^{78}	0.12	0.17	0.09	0.06	0.04	4.68×10^{-1}	1.37×10^{-2}	4.83×10^{-2}	2.44	1.78×10^{2}
K^{82}	0.23	0.41	0.31	0.24	0.11	3.37×10^{-1}	2.17×10^{-3}	1.14×10^{-1}	3.50	1.61×10^{3}
F^{83}	0.17	0.27	0.18	0.17	0.05	2.22×10^{-1}	1.50×10^{-3}	4.91×10^{-2}	3.79	2.53×10^{3}
I^{86}	0.29	0.55	0.52	0.5	0.39	1.67×10^{-1}	8.90×10^{-4}	3.89×10^{-1}	1.55	1.75×10^{3}
R^{87}	0.17	0.33	0.28	0.27	0.14	1.88×10^{-1}	7.60×10^{-4}	1.38×10^{-1}	3.66	4.82×10^{3}
S^{101}	0.11	0.17	0.1	0.09	0.06	6.39×10^{-1}	1.87×10^{-2}	7.17×10^{-2}	1.99	1.07×10^{2}
L^{102}	0.06	0.11	0.08	0.06	0.02	9.96×10^{-2}	2.20×10^{-3}	2.47×10^{-2}	1.55	7.06×10^{2}
C^{151}	0.05	0.06	0.05	0.05	0.04	6.15×10^{-1}	3.75×10^{-2}	4.45×10^{-2}	6.89	1.84×10^{2}
A^{152}	0.01	0.14	0.02	0.02	0.02	1.38×10^{-1}	6.90×10^{-3}	2.10×10^{-2}	8.07	1.17×10^{3}
F^{157}	0.15	0.23	0.19	0.16	0.12	1.38×10^{-1}	2.83×10^{-3}	1.20×10^{-1}	2.55	9.02×10^{2}

†Initial time point excluded from calculations

Table E.4: **Inhibited HRV14-3C K_{ex} and P_{factor} Calculations**

Res	Volumes at Timepoint t_{min}						$V = V_0^{-K_{ex}t} + Baseline$			K_{rc} s^{-1}	P_{factor}
	120	270	420	1340	2885	8822	V_0	K_{ex} s^{-1}	Baseline		(K_{rc} / K_{ex})
F 6	7.64	2.66	1.69	1.05	1.02	1.11	6.54	9.15×10^{-3}	1.09	2.79	3.05×10^{2}
S 9	0.17	0.04	0.04	0.00	0.00	0.00	1.89×10^{1}	5.27×10^{-3}	0.00	5.20	9.86×10^{2}
L 11	1.68	0.25	0.25	0.26	0.43	0.18	1.45	2.75×10^{-2}	2.27×10^{-1}	1.78	6.47×10^{1}
I 15	19.74	18.63	17.14	13.69	8.54	3.98	1.60×10^{1}	4.00×10^{-4}	3.46	2.73	6.83×10^{3}
M 16	14.31	15.24	15.43	15.42	14.9	15.57	0.00	$<1.90 \times 10^{-6}$	0.00	3.09	1.63×10^{6}
I 18	12.82	13.28	12.72	12.67	12.65	12.27	0.00	$<1.90 \times 10^{-6}$	0.00	1.68	8.85×10^{5}
T 19	18.87	19.97	19.66	18.69	18.11	16.95	3.05	2.20×10^{-4}	1.65×10^{1}	3.09	1.41×10^{4}
F 25	11.45	11	11.14	11.62	11.13	11.00	0.00	$<1.90 \times 10^{-6}$	0.00	2.79	$>1.00 \times 10^{6}$
T 26	11.29	11.02	10.46	9.16	6.85	5.37	6.18	4.20×10^{-4}	5.15	4.29	1.02×10^{4}
G 27	8.54	9.11	9.17	8.97	8.71	9.05	0.00	$<1.90 \times 10^{-6}$	0.00	5.20	$>1.00 \times 10^{6}$
L 28	16.80	16.96	17.57	17.05	17.03	16.70	0.00	$<1.90 \times 10^{-6}$	0.00	2.73	$>1.00 \times 10^{6}$
G 29	12.86	12.4	12.86	12.48	12.7	12.67	0.00	$<1.90 \times 10^{-6}$	0.00	4.64	$>1.00 \times 10^{6}$
I 30	9.47	9.34	9.30	9.64	9.67	9.95	0.00	$<1.90 \times 10^{-6}$	0.00	2.31	$>1.00 \times 10^{6}$
H 31	13.98	14.3	14.52	14.3	13.89	13.66	0.00	$<1.90 \times 10^{-6}$	0.00	2.99	$>1.00 \times 10^{6}$
V 34	17.62	18.09	17.97	17.76	17.64	17.81	0.00	$<1.90 \times 10^{-6}$	0.00	2.52	$>1.00 \times 10^{6}$
C 35	10.92	11.13	11.24	11.31	11.33	11.44	0.00	$<1.90 \times 10^{-6}$	0.00	7.45	$>1.00 \times 10^{6}$
V 36	11.63	11.35	11.63	12.32	11.99	12.23	0.00	$<1.90 \times 10^{-6}$	0.00	3.66	$>1.00 \times 10^{6}$
I 37	20.14	20.23	20.69	20.19	20.43	20.48	0.00	$<1.90 \times 10^{-6}$	0.00	1.63	$>1.00 \times 10^{6}$
T 39	15.61	16.57	16.59	16.04	15.96	16.7	0.00	$<1.90 \times 10^{-6}$	0.00	3.06	$>1.00 \times 10^{6}$
A 41	8.20	6.61	5.73	3.69	2.52	1.05	6.24	8.30×10^{-4}	1.38	5.08	6.12×10^{3}
Q 42	7.44	2.90	1.49	0.38	0.56	0.68	6.88	6.99×10^{-3}	5.53×10^{-1}	4.64	6.64×10^{2}
V 47	9.67	5.16	3.05	0.11	0.22	0.33	9.43	4.15×10^{-3}	2.11×10^{-1}	1.61	3.87×10^{2}
L 48	14.54	14.28	13.74	13.23	12.03	9.41	6.56	1.60×10^{-4}	7.80	1.93	1.20×10^{4}
V 49	15.33	15.37	15.76	16.34	16.28	16.21	0.00	$<1.90 \times 10^{-6}$	0.00	1.55	$>1.00 \times 10^{6}$
Q 52	22.42	18.20	14.63	5.48	1.18	1.01	2.12×10^{1}	1.36×10^{-3}	9.68×10^{-1}	5.62	4.13×10^{3}
I 54	17.41	17.99	18.25	17.81	18.27	17.74	0.00	$<1.90 \times 10^{-6}$	0.00	2.18	$>1.00 \times 10^{6}$
L 62	14.57	13.1	11.90	6.94	3.47	0.90	1.34×10^{1}	6.50×10^{-4}	9.61×10^{-1}	2.58	3.97×10^{3}
E 71	14.00	7.44	4.67	0.45	0.61	0.70	1.33×10^{1}	4.18×10^{-3}	5.83×10^{-1}	1.93	4.61×10^{2}
L 72	21.22	20.31	19.64	15.92	11.05	5.81	1.63×10^{1}	3.40×10^{-4}	4.95	1.90	5.60×10^{3}
T 73	22.04	22.39	23.14	22.88	22.87	23.88	0.00	$<1.90 \times 10^{-6}$	0.00	3.16	$>1.00 \times 10^{6}$
V 74	16.64	16.21	16.92	17.18	16.89	17.05	0.00	$<1.90 \times 10^{-6}$	0.00	1.74	$>1.00 \times 10^{6}$
L 75	19.93	19.88	20.34	20.51	20.61	20.55	0.00	$<1.90 \times 10^{-6}$	0.00	1.93	$>1.00 \times 10^{6}$
T 76	18.02	18.07	18.64	17.94	17.12	16.49	1.86×10^{1}	1.00×10^{-5}	4.62×10^{-1}	3.16	3.16×10^{5}
L 77	12.89	12.5	12.27	10.39	9.16	5.91	7.68	2.50×10^{-4}	5.11	1.99	7.97×10^{3}
I 78	6.02	2.07	0.76	0.51	0.74	0.94	5.35	9.56×10^{-3}	6.86×10^{-1}	2.44	2.55×10^{2}

Continued on Next Page...

Table E.4: Inhibited HRV14-3C K_{ex} and P_{factor} Calculations – Continued

Res	Volumes at Timepoint t_{min}						$V = V_0^{-K_{ex}t} + Baseline$			K_{rc}	P_{factor}
	120	270	420	1340	2885	8822	V_0	K_{ex}	Baseline		(K_{rc} / K_{ex})
K 82	16.75	17.09	17.51	17.06	16.9	16.78	0.00	$<1.90 \times 10^{-6}$	0.00	3.50	$>1.00 \times 10^{6}$
F 83	9.68	6.49	4.13	0.61	0.87	0.07	9.34	2.96×10^{-3}	3.82×10^{-1}	3.79	1.28×10^{3}
I 86	12.34	12.19	12.51	12.58	12.53	12.50	0.00	$<1.90 \times 10^{-6}$	0.00	1.55	$>1.00 \times 10^{6}$
R 87	10.38	10.23	10.04	7.74	5.55	3.30	7.60	3.90×10^{-4}	3.01	3.66	9.39×10^{3}
I 90	13.59	9.06	7.64	5.86	6.68	6.10	7.32	6.00×10^{-3}	6.24	2.04	3.40×10^{2}
S 91	6.42	1.22	0.71	0.53	0.09	0.66	5.96	1.34×10^{-2}	4.62×10^{-1}	5.08	3.79×10^{2}
A 99	16.56	16.37	15.53	13.06	9.95	6.45	10.9	3.50×10^{-4}	5.83	3.54	1.01×10^{4}
T 100	15.05	15.34	15.11	15.24	14.87	15.16	0.00	$<1.90 \times 10^{-6}$	0.00	4.01	$>1.00 \times 10^{6}$
L 101	28.55	29.17	28.68	29.66	29.11	29.77	0.00	$<1.90 \times 10^{-6}$	0.00	1.99	$>1.00 \times 10^{6}$
V 102	16.92	14.46	14.78	14.59	14.15	13.55	2.69	1.30×10^{-2}	1.42×10^{1}	1.55	1.20×10^{2}
V 103	12.2	13.16	12.31	12.51	11.99	10.94	13.2	2.00×10^{-5}	5.47×10^{-1}	1.68	8.41×10^{4}
L 113	15.61	15.09	14.04	10.96	7.84	3.57	12.5	3.60×10^{-4}	3.06	1.74	4.83×10^{3}
V 115	17.56	18.27	18.14	18.14	17.21	17.26	9.44	2.70×10^{-4}	1.71×10^{1}	1.66	6.16×10^{3}
G 116	4.07	1.32	0.44	0.45	0.61	0.25	3.67	9.82×10^{-3}	4.05×10^{-1}	5.02	5.11×10^{2}
V 118	10.57	9.74	9.53	8.22	5.63	3.19	7.67	3.20×10^{-4}	2.69	1.50	4.69×10^{3}
A 121	4.33	1.65	0.33	0.33	0.05	0.64	4.10	8.09×10^{-3}	2.59×10^{-1}	4.91	6.07×10^{2}
I 124	6.35	4.59	3.04	0.92	0.10	0.29	6.10	2.33×10^{-3}	2.13×10^{-1}	1.50	6.45×10^{2}
L 126	1.16	0.32	0.07	0.06	0.12	0.05	1.19	9.55×10^{-3}	1.95×10^{-2}	3.23	3.39×10^{2}
T 131	7.45	4.72	2.89	0.97	0.85	0.60	6.70	3.69×10^{-3}	7.73×10^{-1}	3.06	8.29×10^{2}
N 132	6.03	3.14	0.96	0.90	0.05	0.56	5.65	5.91×10^{-3}	4.49×10^{-1}	6.66	1.13×10^{3}
R 133	3.67	0.91	0.34	-0.26	0.24	0.35	3.55	9.85×10^{-3}	1.18×10^{-1}	6.81	6.91×10^{2}
M 134	16.66	17.26	17.29	16.27	15.17	14.15	3.35	2.90×10^{-4}	1.38×10^{1}	5.50	1.90×10^{4}
I 135	11.1	10.96	10.61	10.96	10.56	10.66	0.00	$<1.90 \times 10^{-6}$	0.00	2.16	$>1.00 \times 10^{6}$
Y 137	17.22	18.1	17.94	17.71	17.69	17.31	0.00	$<1.90 \times 10^{-6}$	0.00	4.10	$>1.00 \times 10^{6}$
Y 139	14.13	14.69	14.66	14.89	14.43	14.04	0.00	$<1.90 \times 10^{-6}$	0.00	2.61	$>1.00 \times 10^{6}$
T 141	5.23	1.68	1.56	0.66	0.48	0.91	4.43	8.90×10^{-3}	7.71×10^{-1}	4.01	4.50×10^{2}
Q 145	1.89	0.78	0.96	0.92	0.93	0.98	0.00	$<1.90 \times 10^{-6}$	0.00	5.62	$>1.00 \times 10^{6}$
G 147	4.95	5.42	5.06	4.82	5.05	4.87	0.00	$<1.90 \times 10^{-6}$	0.00	10.9	$>1.00 \times 10^{6}$
G 148	13.9	14.04	14.29	13.87	14.06	13.89	0.00	$<1.90 \times 10^{-6}$	0.00	7.13	$>1.00 \times 10^{6}$
V 149	20.31	20.29	20.03	20.18	20.14	20.55	0.00	$<1.90 \times 10^{-6}$	0.00	2.39	$>1.00 \times 10^{6}$
L 150	12.45	13.32	13.22	12.61	12.51	12.81	0.00	$<1.90 \times 10^{-6}$	0.00	1.93	$>1.00 \times 10^{6}$
C 151	17.35	17.64	17.67	17.99	17.78	17.6	0.00	$<1.90 \times 10^{-6}$	0.00	6.89	$>1.00 \times 10^{6}$
A 152	13.57	14.15	13.26	13.33	13.05	12.34	1.66	2.00×10^{-4}	1.20×10^{1}	8.07	4.03×10^{4}
K 155	7.69	3.11	1.35	0.59	0.32	0.08	7.39	6.47×10^{-3}	3.00×10^{-1}	5.02	$>1.00 \times 10^{6}$
F 157	13.91	13.69	13.43	13.73	13.18	13.46	0.00	$<1.90 \times 10^{-6}$	0.00	2.55	$>1.00 \times 10^{6}$

Continued on Next Page...

Table E.4: Inhibited HRV14-3C K_{ex} and P_{factor} Calculations – Continued

Res	Volumes at Timepoint t_{min}						$V = V_0^{-K_{ex}t} + Baseline$			K_{rc}	P_{factor}
	120	270	420	1340	2885	8822	V_0	K_{ex}	Baseline		(K_{rc} / K_{ex})
G 158	16.98	16.47	16.79	16.7	16.55	16.91	0.00	$<1.90 \times 10^{-6}$	0.00	6.29	$>1.00 \times 10^{6}$
I 159	12.19	12.53	12.3	12.6	12.32	12.35	0.00	$<1.90 \times 10^{-6}$	0.00	2.31	$>1.00 \times 10^{6}$
H 160	10.05	9.89	10.18	9.94	10.44	10.38	0.00	$<1.90 \times 10^{-6}$	0.00	2.99	$>1.00 \times 10^{6}$
V 161	24.23	23.55	23.81	21.93	20.02	16.82	8.26	2.50×10^{-4}	1.59×10^{1}	2.31	9.22×10^{3}
G 162	14.55	15.51	15.01	15.45	14.57	15.07	0.00	$<1.90 \times 10^{-6}$	0.00	5.02	$>1.00 \times 10^{6}$
G 163	12.86	13.2	13.11	13.78	14.03	14.56	0.00	$<1.90 \times 10^{-6}$	0.00	7.13	$>1.00 \times 10^{6}$
N 164	0.68	0.17	0.14	0.08	0.03	0.12	0.66	8.10×10^{-3}	1.45×10^{-2}	9.13	1.13×10^{3}
Q 167	10.45	4.42	2.73	0.75	0.28	0.62	9.79	5.71×10^{-3}	5.88×10^{-1}	5.95	1.04×10^{3}
G 168	14.75	15.35	15.35	14.58	14.62	14.86	0.00	$<1.90 \times 10^{-6}$	0.00	7.37	$>1.00 \times 10^{6}$
F 169	31.76	32.09	32.38	31.82	31.32	31.05	0.00	$<1.90 \times 10^{-6}$	0.00	4.01	$>1.00 \times 10^{6}$
S 170	9.24	8.97	8.99	9.67	9.13	9.36	0.00	$<1.90 \times 10^{-6}$	0.00	7.05	$>1.00 \times 10^{6}$
A 171	10.51	10.77	10.19	10.13	10.51	10.92	0.00	$<1.90 \times 10^{-6}$	0.00	6.08	$>1.00 \times 10^{6}$
Q 172	12.08	12.56	12.43	12.22	12.4	12.05	0.00	$<1.90 \times 10^{-6}$	0.00	5.82	$>1.00 \times 10^{6}$
L 173	13.92	14.38	13.76	13.16	13.08	12.93	0.00	$<1.90 \times 10^{-6}$	0.00	2.83	$>1.00 \times 10^{6}$
K 174	10.73	9.95	10.89	9.17	7.27	5.38	5.89	3.00×10^{-4}	4.89	3.27	1.09×10^{4}